

1	1515	100.0	2190	14	US-10-135-689-1	Sequence 1, Appl
2	1515	100.0	2190	17	US-10-690-617-1	Sequence 1, Appl
3	1510.2	99.7	1937	13	US-10-333-314-39	Sequence 39, Appl
4	1507	99.5	2018	13	US-10-302-172-433	Sequence 43, Appl
5	1505.4	99.4	1611	13	US-10-032-900A-55	Sequence 55, Appl
6	1490.6	98.4	1683	9	US-09-922-138-9	Sequence 9, Appl
7	1490.6	98.4	1686	17	US-10-391-364-43	Sequence 43, Appl
8	1490.6	98.4	2111	9	US-09-922-138-7	Sequence 7, Appl
9	1490.6	98.4	2711	17	US-10-391-364-41	Sequence 41, Appl
10	1457.2	96.2	3501	13	US-10-168-582-14	Sequence 14, Appl
11	1455.8	96.1	1547	13	US-10-092-900A-53	Sequence 53, Appl
12	1381.4	91.2	1725	13	US-10-092-900A-57	Sequence 57, Appl
13	651.2	43.0	4427	15	US-10-316-124-7	Sequence 7, Appl
14	651.2	43.0	4427	15	US-10-316-124-8	Sequence 8, Appl

Db 293 GGTGTGACCCACACAGGGGACAGAGTGCCTCTGTGATCCCTGCGCACTACTTCAAGA 352
QY 181 CTGCTCCAGCCCGGCCCTAGCCTCTACAGCAGGAAGCTTCCCTACAGGAGCGGCCAGCA 240
Db 353 CTGCTCCAGCCCGGCCCTAGCCTCTACAGCAGGAAGCTTCCCTACAGGAGCGGCCAGCA 412
QY 241 GGAAGCTATCTGAGAGCGCAGGCTGGGCCCTTATGCGACGGGGCTGCCAGCCACATCTCC 300
Db 413 GGAAGCTATCTGAGAGCGCAGGCTGGGCCCTTATGCGACGGGGCTGCCAGCCACATCTCC 472
QY 301 CCCCGGCCCTGGCGAGGCCACCATCGAGTCCCAACCACTGCGCATCTCAGATGAGAG 360
Db 473 CCCCGGCCCTGGCGAGGCCACCATCGAGTCCCAACCACTGCGCATCTCAGATGAGAG 532
QY 361 GACTGGCTGAGCTGAACCAAGTCAAGCTGCAGAGTGAGATGGCAAGGGTGGCTACGGT 420
Db 533 GACTGGCTGAGCTGAACCAAGTCAAGCTGCAGAGTGAGATGGCAAGGGTGGCTACGGT 592
QY 421 GTGCTGAGGCTGGCTTACAGCAAGTGAAGAGACAGACACTATGCAATGAAGTCTTTTCC 480
Db 593 GTGCTGAGGCTGGCTTACAGCAAGTGAAGAGACAGACACTATGCAATGAAGTCTTTTCC 652
QY 481 AAAAAGAAGTTACTGAAGCAGTATGCTTTTCCAGTGGCCCTCCCGAGAGGGTCCAG 540
Db 653 AAAAAGAAGTTACTGAAGCAGTATGCTTTTCCAGTGGCCCTCCCGAGAGGGTCCAG 712
QY 541 GCTGCCAGGAGGACCAAGCAGAGCTGCTGCCCTGGAGCGGGTGTACCAAGAGATT 600
Db 713 GCTGCCAGGAGGACCAAGCAGAGCTGCTGCCCTGGAGCGGGTGTACCAAGAGATT 772
QY 601 GCCATCTGAAGAGCTGGACAGCTGAATGTGGTCAAACTGATCGAGTCTCTGGATGAC 660
Db 773 GCCATCTGAAGAGCTGGACAGCTGAATGTGGTCAAACTGATCGAGTCTCTGGATGAC 832
QY 661 CCAAGCTGAGGACCACTTATTTGGTGTGTGACTCTGAGAAAGGGGCCCTCATGGAA 720
Db 833 CCAAGCTGAGGACCACTTATTTGGTGTGTGACTCTGAGAAAGGGGCCCTCATGGAA 892
QY 721 GTGCTCTGACAAAGCCCTTCTCGAGGAGCAAGCTGGCTCTACCTGCGGGACGTATC 780
Db 893 GTGCTCTGACAAAGCCCTTCTCGAGGAGCAAGCTGGCTCTACCTGCGGGACGTATC 952
QY 781 CTGGGCTCGAGTACTGCTCCAGAGATCGTCCAGAGCATCGTCCAGAGCATCAGCCATCCAC 840
Db 953 CTGGGCTCGAGTACTGCTCCAGAGATCGTCCAGAGCATCGTCCAGAGCATCAGCCATCCAC 1012
QY 841 CTGCTCTGGGGATGATGGGACGTGAAGATCGCGACTTTTGGCGTCAAGCAACCAAGTTT 900
Db 1013 CTGCTCTGGGGATGATGGGACGTGAAGATCGCGACTTTTGGCGTCAAGCAACCAAGTTT 1072
QY 901 GAGGGAAACGACGCTCAGCTGTCAGACGCGGGGAAACCCAGCATTCATGGCCCCCGAG 960
Db 1073 GAGGGAAACGACGCTCAGCTGTCAGACGCGGGGAAACCCAGCATTCATGGCCCCCGAG 1132
QY 961 GCATTTCTGATTCGGCCAGAGCTTCAGTGGGAAGCCCTTGATGATGGGCCACTGGC 1020
Db 1133 GCATTTCTGATTCGGCCAGAGCTTCAGTGGGAAGCCCTTGATGATGGGCCACTGGC 1192
QY 1021 GTACGTTTCTATCTGTTTCTATGGAAGTGCCTTATTCATGACGATTCATCTCTGGCC 1080
Db 1193 GTACGTTTCTATCTGTTTCTATGGAAGTGCCTTATTCATGACGATTCATCTCTGGCC 1252
QY 1081 CTCCACAGGAAGTCAAGATGAGCCGCTGGTGTTCCTGAGGAGCCAGAAATCAGCGAG 1140
Db 1253 CTCCACAGGAAGTCAAGATGAGCCGCTGGTGTTCCTGAGGAGCCAGAAATCAGCGAG 1312
QY 1141 GAGCTCAAGGACCTGATCTGAAGATGTTAGACAAGATCCCGAGACGAGATTGGGGTG 1200
Db 1313 GAGCTCAAGGACCTGATCTGAAGATGTTAGACAAGATCCCGAGACGAGATTGGGGTG 1372
QY 1201 CCAGACATCAAGTTGACCTTGGGTGACAGAAACGGGAGAGGCCCTTCTCTCGGAG 1260

Db 1373 CCAGACATCAAGTTGCAACCTTGGGTGACCAAGAAACGGGAGAGCCCTTCTCTCGGAG 1432
QY 1261 GAGGAGCACTGCAGCGTGGTGGAGGTGACAGAGGGGGAGGTAAAGAACTCAGTCAGGCTC 1320
Db 1433 GAGGAGCACTGCAGCGTGGTGGAGGTGACAGAGGGGGAGGTAAAGAACTCAGTCAGGCTC 1492
QY 1321 ATCCCAAGCTGACCAAGCTGATCCCTGGTGAAGTCCATGCTGAGGAAGCTTCTTTGGG 1380
Db 1493 ATCCCAAGCTGACCAAGCTGATCCCTGGTGAAGTCCATGCTGAGGAAGCTTCTTTGGG 1552
QY 1381 AACCCGTTTGGAGCCCGAGGCAAGGAGGAGAGGATCCATGCTCTCCAGGAAACCTA 1440
Db 1553 AACCCGTTTGGAGCCCGAGGCAAGGAGGAGAGGATCCATGCTCTCCAGGAAACCTA 1612
QY 1441 CTGCTGAAAGAGAGGGTTTGGTGAAGGGGCAAGAGCCCAAGAGTCCCGGCGTCCAGGAA 1500
Db 1613 CTGCTGAAAGAGAGGGTTTGGTGAAGGGGCAAGAGCCCAAGAGTCCCGGCGTCCAGGAA 1672
QY 1501 GACGAGGCTGCATCC 1515
Db 1673 GACGAGGCTGCATCC 1687
RESULT 2
US-10-690-617-1
; Sequence 1, Application US/10690617
; Publication No. US20040086926A1
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL000904DIV II
; CURRENT APPLICATION NUMBER: US/10/690,617
; CURRENT FILING DATE: 2003-10-23
; PRIOR APPLICATION NUMBER: 60/247,031
; PRIOR FILING DATE: 2000-11-13
; PRIOR APPLICATION NUMBER: 09/729,995
; PRIOR FILING DATE: 2000-12-06
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2190
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-690-617-1
Query Match 100.0%; Score 1515; DB 17; Length 2190;
Best Local Similarity 100.0%; Pred No. 0;
Matches 1515; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGAGGGGGTCCAGCTCTGCTGCCAGGATCTCTCGGCGAGAGCTGTGAGAACGGGTG 60
Db 173 ATGAGGGGGTCCAGCTCTGCTGCCAGGATCTCTCGGCGAGAGCTGTGAGAACGGGTG 232
QY 61 GCAGCCATCGATGTGACTCACTTGGAGGAGGAGATGTTGGCCCAAGCTTACTAGAAC 120
Db 233 GCAGCCATCGATGTGACTCACTTGGAGGAGGAGATGTTGGCCCAAGCTTACTAGAAC 292
QY 121 GGTGTGGAACCCCGAGGCGGAGAGCTGCTCTGTGATCCCTGGCAGTACTTCAAGA 180
Db 293 GGTGTGGAACCCCGAGGCGGAGAGCTGCTCTGTGATCCCTGGCAGTACTTCAAGA 352
QY 181 CTGCTCCAGCGCGGCTTAGCCTCTACGCAAGAGCTTTCCCTACAGAGCGGCCAGCA 240
Db 353 CTGCTCCAGCGCGGCTTAGCCTCTACGCAAGAGCTTTCCCTACAGAGCGGCCAGCA 412
QY 241 GGAAGCTATCTGAGGCGCAGGCTGGGCCCTTATGCCACGGGCTGCCAGGCCACATCTCC 300
Db 413 GGAAGCTATCTGAGGCGCAGGCTGGGCCCTTATGCCACGGGCTGCCAGGCCACATCTCC 472
QY 301 CCCCGGCCCTGGCGAGGCCACCATCGAGTCCCAACCACTGCGCATCTCAGATGAGAG 360

Db 473 CCCCAGGCTGGCGAGGCCACCATCGAGTCCACCACTGGCCATCTCAGATGCAGAG 532
QY 361 GACTGGTGCAGCTGAACACAGTACAGCTGCAGAGTGCAGATGGCAAGGGTGCCTACGGT 420
Db 533 GACTGGTGCAGCTGAACACAGTACAGCTGCAGAGTGCAGATGGCAAGGGTGCCTACGGT 592
QY 421 GTGGTGAAGGCTGCTTACACAGAAAGTGAAGACAGACACTATCAATGAAGTCCCTTTCC 480
Db 593 GTGGTGAAGGCTGCTTACACAGAAAGTGAAGACAGACACTATCAATGAAGTCCCTTTCC 652
QY 481 AAAAAGAGTTACTGAAGCAGTAGTGTCTTCCACGTGCGCCCTCCCGAGAGGGTCCCGAG 540
Db 653 AAAAAGAGTTACTGAAGCAGTAGTGTCTTCCACGTGCGCCCTCCCGAGAGGGTCCCGAG 712
QY 541 GCTGCCAGGAGGACACAGCAGCAGCTGCTGCCCTCGAGCGGGTGTACACAGAGATT 600
Db 713 GCTGCCAGGAGGACACAGCAGCAGCTGCTGCCCTCGAGCGGGTGTACACAGAGATT 772
QY 601 GCCATCTGAAGAGCTGGACCACTGAATGTGTCAAACTGATCGAGGTCTCTGGATGAC 660
Db 773 GCCATCTGAAGAGCTGGACCACTGAATGTGTCAAACTGATCGAGGTCTCTGGATGAC 832
QY 661 CCAGCTGAGGACAACTTATTTGGTGTGTTGACCTCTCTGAGAAAGGGGCCCGTCAATGGAA 720
Db 833 CCAGCTGAGGACAACTTATTTGGTGTGTTGACCTCTCTGAGAAAGGGGCCCGTCAATGGAA 892
QY 721 GTGCCCTGACAGCCCTTCTCGGAGGACAGCTGCGCTCTACCTCGCGGACGTCTATC 780
Db 893 GTGCCCTGACAGCCCTTCTCGGAGGACAGCTGCGCTCTACCTCGCGGACGTCTATC 952
QY 781 CTGGCCCTCAGTACTTGCATCGCAGAAAGATCTCCACAGGACATCAAGCCATCCAAAC 840
Db 953 CTGGCCCTCAGTACTTGCATCGCAGAAAGATCTCCACAGGACATCAAGCCATCCAAAC 1012
QY 841 CTCCTCTCGGGGATGATGGCAGCTGAAGATCCCGCACTTTGGCGTCAAGAACAGTTT 900
Db 1013 CTCCTCTCGGGGATGATGGCAGCTGAAGATCCCGCACTTTGGCGTCAAGAACAGTTT 1072
QY 901 GAGGGACACAGCTCAGCTGTCCAGCAGCGCGGMAACCCAGCATTCATGGCCCCCGAG 960
Db 1073 GAGGGACACAGCTCAGCTGTCCAGCAGCGCGGMAACCCAGCATTCATGGCCCCCGAG 1132
QY 961 GCATTTCTGATTCGGCCAGAGCTTCAGTGGGAAGCCTTGGATGTATGGGCCACTGGC 1020
Db 1133 GCATTTCTGATTCGGCCAGAGCTTCAGTGGGAAGCCTTGGATGTATGGGCCACTGGC 1192
QY 1021 GTCAGTTGTACTCTTGTCTATGGGAAGTGCACATTCATCGACGATTTTCATCTGGCC 1080
Db 1193 GTCAGTTGTACTCTTGTCTATGGGAAGTGCACATTCATCGACGATTTTCATCTGGCC 1252
QY 1081 CTCACAGGAGATCAAGATGAGCCGTGTGTCTCTGAGGAGCCAGAAATCAGCGAG 1140
Db 1253 CTCACAGGAGATCAAGATGAGCCGTGTGTCTCTGAGGAGCCAGAAATCAGCGAG 1312
QY 1141 GAGCTCAAGGACCTGATCCTGAAGATGTAGACAAAGATCCCGAGACAGAAATTTGGGGTG 1200
Db 1313 GAGCTCAAGGACCTGATCCTGAAGATGTAGACAAAGATCCCGAGACAGAAATTTGGGGTG 1372
QY 1201 CCAGACATCAAGTTGTACCCCTTGGGTGACCAAGAACGGGGAGGAGCCCTTCTCTCGGAG 1260
Db 1373 CCAGACATCAAGTTGTACCCCTTGGGTGACCAAGAACGGGGAGGAGCCCTTCTCTCGGAG 1432
QY 1261 GAGGAGCACTCAGCGTGTGAGGTGACAGAGGGGAGGTTAAGAACTCAGTCAGGCTC 1320
Db 1433 GAGGAGCACTCAGCGTGTGAGGTGACAGAGGGGAGGTTAAGAACTCAGTCAGGCTC 1492
QY 1321 ATCCCACTGACACACCGTGTATCTCTGTGAGTCCATGCTCAGGAAGCGTTCTCTTTGGG 1380
Db 1493 ATCCCACTGACACACCGTGTATCTCTGTGAGTCCATGCTCAGGAAGCGTTCTCTTTGGG 1552
QY 1381 AACCCGTTTGAAGCCCAAGGCAAGGAGGAAAGAGCGATCCATGTCTGTCTCCAGGAACCTTA 1440
Db 1553 AACCCGTTTGAAGCCCAAGGCAAGGAGGAAAGAGCGATCCATGTCTGTCTCCAGGAACCTTA 1612

QY 1441 CTGGTGAAGAAAGGGTTTGGTGAAGGGGCAAGAGCCCAAGAGCTCCCGCGGTCCAGGAA 1500
Db 1613 CTGGTGAAGAAAGGGTTTGGTGAAGGGGCAAGAGCCCAAGAGCTCCCGCGGTCCAGGAA 1672
QY 1501 GACGAGGCTGCATCC 1515
Db 1673 GACGAGGCTGCATCC 1687

RESULT 3

US-10-333-314-39
; Sequence 39, Application US/10333314
; Publication No. US20030211093A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.; YUE, Henry
; APPLICANT: KHAN, Farrah A.; GURUPAJAN, Rajagopal
; APPLICANT: HAFALIA, Chandra S.; CHAWLA, Narinder K.
; APPLICANT: ARVIZU, Ameena R.; RAMKUMAR, Jayalaxmi
; APPLICANT: GANDHI, Mariah R.; POLICKY, Jennifer L.
; APPLICANT: BAUGHN, Michael B.; TRIBOULEY, Catherine M.
; APPLICANT: THORNTON, Michael B.; BANDMAN, Olga
; APPLICANT: NGUYEN, Daniel B.; LU, Yan
; APPLICANT: BURFORD, Neil; LAL, Preeti G.
; APPLICANT: DING, Li; YAO, Monique G.
; APPLICANT: ELLIOTT, Vicki S.; RECIPON, Shirley A.
; APPLICANT: KEARNEY, Liam; LU, Dying Aina M.
; APPLICANT: GREENWALD, Sara R.; TANG, Y. Tom
; APPLICANT: XU, Yuming; WALSH, Roderick T.
; APPLICANT: GIETZEN, Kimberly J.; YANG, Junming
; APPLICANT: JACKSON, Jennifer L.
; TITLE OF INVENTION: HUMAN KINASES
; FILE REFERENCE: PI-0162 USN
; CURRENT APPLICATION NUMBER: US/10/333,314
; CURRENT FILING DATE: 2003-01-15
; PRIOR APPLICATION NUMBER: PCT/US01/23092
; PRIOR FILING DATE: 2001-07-20
; PRIOR APPLICATION NUMBER: US 60/220,038
; PRIOR FILING DATE: 2000-07-21
; PRIOR APPLICATION NUMBER: US 60/222,112
; PRIOR FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: US 60/222,831
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: US 60/224,729
; PRIOR FILING DATE: 2000-08-11
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: PERL Program
; SEQ ID NO 39
; LENGTH: 1937
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: 2190612CB1
US-10-333-314-39

Query Match 99.7%; Score 1510.2; DB 13; Length 1937;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1512; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ATGAGAGGGGTCCAGTGTCTGTCGCCAGGATCCTCGGCGAGAGCTGTTAGAACGGGTG 60
Db 65 ATGAGAGGGGTCCAGTGTCTGTCGCCAGGATCCTCGGCGAGAGCTGTTAGAACGGGTG 124
QY 61 GCACCCATCATGTGATCTCACTTGGAGAGGCGACATGTCGCCCCAGAGCCTTACTAGAAAC 120
Db 125 GCACCCATCATGTGATCTCACTTGGAGAGGCGATGTTGCCCCAGAGCCTTACTAGAAAC 184
QY 121 GGTGTGACACCCCAACACCGGGCCAGAGCTGCTCTGTGATCCCTGGCAGTACTTCAAGA 180
Db 185 GGTGTGACACCCCAACACCGGGCCAGAGCTGCTCTGTGATCCCTGGCAGTACTTCAAGA 244
QY 181 CTGCTCCAGCGCCCGCTAGCCTCTCAGCCAGGAAGCTTTCCTCTACAGGAGCGGCCGCA 240

Db 245 CTGCTCCAGCCGCGCTAGCCTCTCAGCCAGAGAGCTTTCCCTACAGGAGCGGCCAGCA 304
QY 241 GGAAGCTATCTGAGGCGCAGGTGGGCCCTTATGCGACGGGGCTGCCAGCCACATCTCC 300
Db 305 GGAAGCTATCTGAGGCGCAGGTGGGCCCTTATGCGACGGGGCTGCCAGCCACATCTCC 364
QY 301 CCCCGGCTGGCGGAGGCCACCATCGAGTCCACACGCTGGCCATCTCAGATGAGAG 360
Db 365 CCCCGGCTGGCGGAGGCCCATCGAGTCCACACGCTGGCCATCTCAGATGAGAG 424
QY 361 GACTGGCTGAGCTGAACCAAGTACAACTGCAGAGTGAGATTGGCAAGGTGCTACGGT 420
Db 425 GACTGGCTGAGCTGAACCAAGTACAACTGCAGAGTGAGATTGGCAAGGTGCTACGGT 484
QY 421 GTGGTGAAGCTGCCCTACACGAAAGTGAAGACAGACATATCAATGAAGTCTTTTC 480
Db 485 GTGGTGAAGCTGCCCTACACGAAAGTGAAGACAGACATATCAATGAAGTCTTTTC 544
QY 481 AAAAAAGATTACTGAAGCAGTATGCTTTCCACGTCGCGCTCCCGGAGGGTCCAG 540
Db 545 AAAAAAGATTACTGAAGCAGTATGCTTTCCACGTCGCGCTCCCGGAGGGTCCAG 604
QY 541 GCTGCCAGGAGGACCCAGCAAGCAGCTGCTGCCCTGGAGCGGTGTACCAAGAGATT 600
Db 605 GCTGCCAGGAGGACCCAGCAAGCAGCTGCTGCCCTGGAGCGGTGTACCAAGAGATT 664
QY 601 GCATCTCTGAAGCTGACACCGTGAATGTGGTCAAACTGATGAGGTCTCTGGATGAC 660
Db 665 GCATCTCTGAAGCTGACACCGTGAATGTGGTCAAACTGATGAGGTCTCTGGATGAC 724
QY 661 CCAGCTGAGCAACCTCTATTGGTGTGTGACCTCTGAGAAAGGGCCGCTCATGAA 720
Db 725 CCAGCTGAGCAACCTCTATTGGTGTGTGACCTCTGAGAAAGGGCCGCTCATGAA 784
QY 721 GTGCCCTGTGAACGCCCTCTCGGAGGACGAGCTCGCTCTACCTGGGGACGCTATC 780
Db 785 GTGCCCTGTGAACGCCCTCTCGGAGGACGAGCTCGCTCTACCTGGGGACGCTATC 844
QY 781 CTGGGCTCGAGTCTTGACCTGCCAGAGATCGTCCAGGGACATCAAGCATCCAC 840
Db 845 CTGGGCTCGAGTCTTGACCTGCCAGAGATCGTCCAGGGACATCAAGCATCCAC 904
QY 841 CTGCTCCTGGGGGATGATGGGCAGGTGAAGATCGCGACTTTGGCGTCAGCAACCATGTT 900
Db 905 CTGCTCCTGGGGGATGATGGGCAGGTGAAGATCGCGACTTTGGCGTCAGCAACCATGTT 964
QY 901 GAGGGAAACGACGCTCAGCTGTCCAGCACGGCGGGAACCCAGCATTCATGGCCCCCGAG 960
Db 965 GAGGGAAACGACGCTCAGCTGTCCAGCACGGCGGGAACCCAGCATTCATGGCCCCCGAG 1024
QY 961 GCCATTTCTGATTCGGCCAGAGCTTCAGTGGGAAGCCCTTGATGTATGGGCCACTGGC 1020
Db 1025 GCCATTTCTGATTCGGCCAGAGCTTCAGTGGGAAGCCCTTGATGTATGGGCCACTGGC 1084
QY 1021 GTACGTTGTACTGTCTTGTATGGGAAGTCCCATTCATCGACGATTTTCATCTGGCC 1080
Db 1085 GTACGTTGTACTGTCTTGTATGGGAAGTCCCATTCATCGACGATTTTCATCTGGCC 1144
QY 1081 CTCACAGAGATCAAGATGAGCCCGTGGTGTTCCTGAGGACCCAGAAATCAGCGAG 1140
Db 1145 CTCACAGAGATCAAGAAATGAGCCCGTGGTGTTCCTGAGGACCCAGAAATCAGCGAG 1204
QY 1141 GAGCTCAAGGACCTCATCTGAAGATGTAGACAAGATCCCGAGCAGATTGGGGTG 1200
Db 1205 GAGCTCAAGGACCTCATCTGAAGATGTAGACAAGATCCCGAGCAGATTGGGGTG 1264
QY 1201 CCAGACATCAAGTTGCAACCTTTGGTGACCAAGAACCGGGAGAGCCCTTCCTTCGGAG 1260
Db 1265 CCAGACATCAAGTTGCAACCTTTGGTGACCAAGAACCGGGAGAGCCCTTCCTTCGGAG 1324
QY 1261 GAGGAGCACTGACGCTGGTGAAGTGCACAGGGGGAGGTTAAGAACTCAGTCAGGCTC 1320

Db 1325 GAGGAGCACTGCAGCGTGGTGGAGGTGACAGAGGAGAGGTTAAGAACTCAGTCAGGCTC 1384
QY 1321 ATCCCAGCTGGACACACGCTGATCCTGGTGAAGTCCATGCTGAGGAAGCGTTCCTTTGGG 1380
Db 1385 ATCCCAGCTGGACACACGCTGATCCTGGTGAAGTCCATGCTGAGGAAGCGTTCCTTTGGG 1444
QY 1381 AACCCGTTTGAAGCCCGCAGGACCGAGGAAAGAGCGATCCATGTCTGCAGGAAACCTA 1440
Db 1445 AACCCGTTTGAAGCCCGCAGGACCGAGGAAAGAGCGATCCATGTCTGCAGGAAACCTA 1504
QY 1441 CTGCTGAAGAAGGGTTTGTGAAGGGGCAAGAGCCAGAGCTCCCGCGCTCCAGAA 1500
Db 1505 CTGCTGAAGAAGGGTTTGTGAAGGGGCAAGAGCCAGAGCTCCCGCGCTCCAGAA 1564
QY 1501 GACGAGGCTGCATCC 1515
Db 1565 GACGAGGCTGCATCC 1579

RESULT 4

US-10-302-172-433
; Sequence 433, Application US/10302172
; Publication No. US20040053250A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. US20040053250A1 Arginine-rich Protein-like Nucleic Acids ar
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 803 1CNCIP
; CURRENT APPLICATION NUMBER: US/10/302,172
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/225,251
; PRIOR FILING DATE: 2002-08-20
; PRIOR APPLICATION NUMBER: PCT US02/05095
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 09/799,451
; PRIOR FILING DATE: 2001-03-05
; NUMBER OF SEQ ID NOS: 950
; SOFTWARE: pt_FL_genes Version 2.0
; SEQ ID NO 433
; LENGTH: 2018
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1746)
US-10-302-172-433

Query Match 99.5%; Score 1507; DB 13; Length 2018;
Best Local Similarity 99.7%; Pred. No. 0;

Matches 1510; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 ATGAGAGGGGGTCCAGCTGTCTGTGTCAGAGATCTCGGGCAGAGCTGTAGAACGGTG 60
Db 232 ATGAGAGGGGGTCCAGCTGTCTGTGTCAGAGATCTCGGGCAGAGCTGTAGAACGGTG 291
QY 61 GCAGCCATCGATGTGACTCACTTGGAGGAGGAGATGGTGGCCCGAGAGCTACTAGAAC 120
Db 292 GCAGCCATCGATGTGACTCACTTGGAGGAGGAGATGGTGGCCCGAGAGCTACTAGAAC 351
QY 121 GGTGTGACCCCCACACCGCCAGAGCTGCCTCTGTATCCCTGGCAGTACTTCAAGA 180
Db 352 GGTGTGACCCCCACACCGCCAGAGCTGCCTCTGTATCCCTGGCAGTACTTCAAGA 411
QY 181 CTGCTCCAGCCCGGCTTAGCTCTTCAGCCAGAGAGCTTTCCCTACAGAGCGGCAGCA 240
Db 412 CTGCTCCAGCCCGGCTTAGCTCTTCAGCCAGAGAGCTTTCCCTACAGAGCGGCAGCA 471
QY 241 GGAAGCTATCTGAGGCGCAGCTGGGCTTATGCGACGGGCGCTGCAGGCCACATCTCC 300
Db 472 GGAAGCTATCTGAGGCGCAGCTGGGCTTATGCGACGGGCGCTGCAGGCCACATCTCC 531

QY 301 CCCGGGCTGGCGAGGCCACCATCGAGTCCACCACTGGCCATCTCAGATGCAG 360
D 532 CCCGGGCTGGCGAGGCCACCATCGAGTCCACCACTGGCCATCTCAGATGCAG 591
QY 361 GACTGCGTGCAGCTGAACCACTGACAGTGCAGAGTGGAGTGGAGGGTCCAGGT 420
D 592 GACTGCGTGCAGCTGAACCACTGACAGTGCAGAGTGGAGTGGAGGGTCCAGGT 651
QY 421 GTGCTGAGGCTGGCTCAACGGAAGTGAACACAGACACTATGCAATGAAAGTCTTCC 480
D 652 GTGCTGAGGCTGGCTCAACGGAAGTGAACACAGACACTATGCAATGAAAGTCTTCC 711
QY 481 AAAAAGAGTTACTGAAGCAGTATGGCTTTCCAGTGGCCCTCCCGGAGAGGTCCCG 540
D 712 AAAAAGAGTTACTGAAGCAGTATGGCTTTCCAGTGGCCCTCCCGGAGAGGTCCCG 771
QY 541 GCTGCCCGAGGAGGACCAAGCAGTGTGCGCTGGAGCGGTGTACAGAGATT 600
D 772 GCTGCCCGAGGAGGACCAAGCAGTGTGCGCTGGAGCGGTGTACAGAGATT 831
QY 601 GCCATCCTGAAGAAGCTGGACCACTGAACTGTGTCAAACTGATCGAGGTCTCGATGAC 660
D 832 GCCATCCTGAAGAAGCTGGACCACTGAACTGTGTCAAACTGATCGAGGTCTCGATGAC 891
QY 661 CCAGCTGAGGACAACTTATTTGGTGTGACCTCTCTGAAAGGGGCCGTCAATGAA 720
D 892 CCAGCTGAGGACAACTTATTTGGTGTGACCTCTCTGAAAGGGGCCGTCAATGAA 951
QY 721 GTGCGCTGTGACAGCCCTCTCGAGGAGCAAGCTGCGCTTACCTGCGGAGTCAATC 780
D 952 GTGCGCTGTGACAGCCCTCTCGAGGAGCAAGCTGCGCTTACCTGCGGAGTCAATC 1011
QY 781 CTGGCCCTCGAGTACTTCTGCACTGCGAGCAAGTCTGCAAGGAGCATCAAGCCATCAAC 840
D 1012 CTGGCCCTCGAGTACTTCTGCACTGCGAGCAAGTCTGCAAGGAGCATCAAGCCATCAAC 1071
QY 841 CTGCTCTGGGGATGATGGGCACTGGAAGTCCCGACTTTGGCGTCAAGCAACAGTTT 900
D 1072 CTGCTCTGGGGATGATGGGCACTGGAAGTCCCGACTTTGGCGTCAAGCAACAGTTT 1131
QY 901 GAGGGAGCAAGCTGAGCTGTCCAGCAGCGCGGAGCCAGCATTCATGCGCCCGAG 960
D 1132 GAGGGAGCAAGCTGAGCTGTCCAGCAGCGCGGAGCCAGCATTCATGCGCCCGAG 1191
QY 961 GCCATTTCTGATTCGGCCAGAGCTTCAGTGGGAAGGCTTGGATGATGGGCCACTGGC 1020
D 1192 GCCATTTCTGATTCGGCCAGAGCTTCAGTGGGAAGGCTTGGATGATGGGCCACTGGC 1251
QY 1021 GTCAGTTGATCTGCTTTGCTATGGGAAGTCCCATTCATCGAGATTTCATCTGGCC 1080
D 1252 GTCAGTTGATCTGCTTTGCTATGGGAAGTCCCATTCATCGAGATTTCATCTGGCC 1311
QY 1081 CTCACAGGAGATCAAGATGAGCCGTGTGTTCTCGAGGAGCCAGAAATCAGCAG 1140
D 1312 CTCACAGGAGATCAAGATGAGCCGTGTGTTCTCGAGGAGCCAGAAATCAGCAG 1371
QY 1141 GAGCTCAAGGACCTGATCTTGAAGTGTAGACAAGAAATCCCGAGACGAGAATGGGGT 1200
D 1372 GAGCTCAAGGACCTGATCTTGAAGTGTAGACAAGAAATCCCGAGACGAGAATGGGGT 1431
QY 1201 CCAGACATCAAGTTGACCCCTTGGGTGACCAAGAACGGGGAGGAGCCCTTCCTTCGGAG 1260
D 1432 CCAGACATCAAGTTGACCCCTTGGGTGACCAAGAACGGGGAGGAGCCCTTCCTTCGGAG 1491
QY 1261 GAGGACACTGACGGTGTGAGGTGACAGAGGGGAGGTATAGAACTAGTCAGCTC 1320
D 1492 GAGGACACTGACGGTGTGAGGTGACAGAGGGGAGGTATAGAACTAGTCAGCTC 1551
QY 1321 ATCCCCAGCTGGACACCGTGATCTGCTGAAGTCCATGCTGAGGAAGCGTTCTTTGGG 1380
D 1552 ATCCCCAGCTGGACACCGTGATCTGCTGAAGTCCATGCTGAGGAAGCGTTCTTTGGG 1611
QY 1381 AACCCGTTTGAGCCCGAGGACGAGGAGGAGGAGCATTCATGCTGCTCCAGGAACCTA 1440

D 1612 AACCCGTTTGAGCCCGAGGAGGAGGAGGAGCATTCATGCTGCTCCAGGAACCTA 1671
QY 1441 CTGCTGAAGAGAGGTTTGGTGAAGGGGGCAAGAGCCAGAGCTCCCGCGGTCCAGAA 1500
D 1672 CTGCTGAAGAGAGGTTTGGTGAAGGGGGCAAGAGCCAGAGCTCCCGCGGTCCAGAA 1731
QY 1501 GACGAGGCTGCATCC 1515
D 1732 GACGAGGCTGCATCC 1746
RESULT 5
US-10-092-900A-55
; Sequence 55, Application US/10092900A
; Publication No. US20040043382A1
; GENERAL INFORMATION:
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Shenoy, Suresh G.
; APPLICANT: Taupier Jr., Raymond J.
; APPLICANT: Pena, Carol E.A.
; APPLICANT: Li, Li
; APPLICANT: Zerhusen, Bryan D.
; APPLICANT: Gusev, Vladimir Y.
; APPLICANT: Ji, Weizhen
; APPLICANT: Gorman, Linda
; APPLICANT: Miller, Charles B.
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Patturajan, Meera
; APPLICANT: Gangolli, Esha A.
; APPLICANT: Vernet, Corine A.M.
; APPLICANT: Guo, Xiaojia Sasha
; APPLICANT: Tchernev, Velizar T.
; APPLICANT: Fernandes, Elma R.
; APPLICANT: Casman, Stacie J.
; APPLICANT: Malyankar, Uriel M.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Liu, Yi
; APPLICANT: Anderson, David W.
; APPLICANT: Spaderna, Steven K.
; APPLICANT: Catterton, Elina
; APPLICANT: Leite, Mario W.
; APPLICANT: Zhong, Haihong
; APPLICANT: Alsobrook, John P.
; APPLICANT: Lepley, Denise M.
; APPLICANT: Rieger, Daniel K.
; APPLICANT: Burgess, Catherine E.
; TITLE OF INVENTION: No. US20040043382A1el Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-290C
; CURRENT APPLICATION NUMBER: US/10/092,900A
; CURRENT FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: USSN 60/274,322
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: USSN 60/283,675
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: USSN 60/338,092
; PRIOR FILING DATE: 2001-12-03
; PRIOR APPLICATION NUMBER: USSN 60/274,281
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: USSN 60/274,191
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: USSN 60/325,681
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: USSN 60/304,354
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: USSN 60/279,995
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: USSN 60/294,899
; PRIOR FILING DATE: 2001-05-31
; PRIOR APPLICATION NUMBER: USSN 60/287,424
; PRIOR FILING DATE: 2001-04-30
; Remaining Prior Application data removed - See File Wrapper or PALM.

; NUMBER OF SEQ ID NOS: 768
; SEQ ID NO 55
; LENGTH: 1611
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (52)..(1567)
US-10-092-900A-55

Query Match 99.4%; Score 1505.4; DB 13; Length 1611;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1509; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 ATGGAGGGGGTCCAGCTGCTGCTGCCAGGATCCTCGGSCAGAGCTGGTAGAACGGGTG 60
DB 52 ATGGAGGGGGTCCAGCTGCTGCTGCCAGGATCCTCGGSCAGAGCTGGTAGAACGGGTG 111

QY 61 GCAGCCATCGATGTGACTCACTTGGAGGAGGAGATGGTGGCCAGAGCCTACTAGAAAC 120
DB 112 GCAGCCATCGATGTGACTCACTTGGAGGAGGAGATGGTGGCCAGAGCCTACTAGAAAC 171

QY 121 GGTGTGACCCCCACCAACCGGSCAGAGCTGCTCTGTGATCCCTGGGAGTACTTCAAGA 180
DB 172 GGTGTGACCCCCACCAACCGGSCAGAGCTGCTCTGTGATCCCTGGGAGTACTTCAAGA 231

QY 181 CTGCTCCAGCGGGCTAGCCTCTCAGCCAGGAAGCTTCCCTACAGGAGCGGCCAGCA 240
DB 232 CTGCTCCAGCGGGCTAGCCTCTCAGCCAGGAAGCTTCCCTACAGGAGCGGCCAGCA 291

QY 241 GGAAGCTATCTGAGGCGCAGGCTGGGCCCTTATGCCACGGGGCTCCAGCCACATCTCC 300
DB 292 GGAAGCTATCTGAGGCGCAGGCTGGGCCCTTATGCCACGGGGCTCCAGCCACATCTCC 351

QY 301 CCCCGGGCTGCGGAGGCGCCACCATCGATCGATGCCACCAAGTGGCCATCTCAGATGACAG 360
DB 352 CCCCGGGCTGCGGAGGCGCCACCATCGATCGATGCCACCAAGTGGCCATCTCAGATGACAG 411

QY 361 GACTGGGTGAGCTGAACAGTACAAAGCTGCAGAGTGAGATGGCAAGGGTGCTACGGT 420
DB 412 GACTGGGTGAGCTGAACAGTACAAAGCTGCAGAGTGAGATGGCAAGGGTGCTACGGT 471

QY 421 GTGGTGAGGCTGCGCTTACAAACGAAAGTGAAGACAGACATATGCAATGAAGTCTTTC 480
DB 472 GTGGTGAGGCTGCGCTTACAAACGAAAGTGAAGACAGACATATGCAATGAAGTCTTTC 531

QY 481 AAAAAGAGTTACTGAAGCAGTATGCTTTCCACGTGCGCTCCCGGAGAGGTCCTCCAG 540
DB 532 AAAAAGAGTTACTGAAGCAGTATGCTTTCCACGTGCGCTCCCGGAGAGGTCCTCCAG 591

QY 541 GCTGCCAGGAGGACCCAGCAGCTGCTGCCCTGGAGCGGGTGTACCAGGAGATT 600
DB 592 GCTGCCAGGAGGACCCAGCAGCTGCTGCCCTGGAGCGGGTGTACCAGGAGATT 651

QY 601 GCATCCTGAAGAGCTGACACAGTGAATGTGGTCAAACTGATCGAGTCTCTGGATGAC 660
DB 652 GCATCCTGAAGAGCTGACACAGTGAATGTGGTCAAACTGATCGAGTCTCTGGATGAC 711

QY 661 CCAGCTGAGGACAACTTATTTGGTGTGACCTCTCAGAAAGGGCCCGTCTATGGAA 720
DB 712 CCAGCTGAGGACAACTTATTTGGTGTGACCTCTCAGAAAGGGCCCGTCTATGGAA 771

QY 721 GTGCCCTGTGACAAGCCCTTCTCGAGGAGCAAGCTCGCCTTACCTCGGGAGCGTCAATC 780
DB 772 GTGCCCTGTGACAAGCTTCTCGAGGAGCAAGCTCGCCTTACCTCGGGAGCGTCAATC 831

QY 781 CTGGGCTCGAGTACTTGCATCTGCCAGAGATGCTCCAGGAGACATCAAGCCATCCAC 840
DB 832 CTGGGCTCGAGTACTTGCATCTGCCAGAGATGCTCCAGGAGACATCAAGCCATCCAC 891

QY 841 CTGCTCTCTGGGGATGATGGGACGCTGAAGATCGCCGACTTTGGCGTCAAGCAACAGTTT 900
DB 892 CTGCTCTCTGGGGATGATGGGACGCTGAAGATCGCCGACTTTGGCGTCAAGCAACAGTTT 951

QY 901 GAGGGGAACGAGCTCAGCTGTCCAGCAGCGGGGAACCCAGCATTCATGGCCCCCGAG 960
DB 952 GAGGGGAACGAGCTCAGCTGTCCAGCAGCGGGGAACCCAGCATTCATGGCCCCCGAG 1011

QY 961 GCCATTTCTGATTCGGGCCAGAGCTTCAGTGGGAAGCCCTTGGATGATGGGSCACTGGC 1020
DB 1012 GCCATTTCTGATTCGGGCCAGAGCTTCAGTGGGAAGCCCTTGGATGATGGGSCACTGGC 1071

QY 1021 GTACGTTGTACTGCTTTGTCTATGGGAAGTGGCCATTCATCGACGATTCATCTCTGGCC 1080
DB 1072 GTACGCTGTACTGCTTTGTCTATGGGAAGTGGCCCTTCATCGACGATTCATCTCTGGCC 1131

QY 1081 CTCCACAGGAAGATCAAGAATGAGCCCGTGGTGTCTTCTGAGGAGCCAGAAATCAGCGAG 1140
DB 1132 CTCCACAGGAAGATCAAGAATGAGCCCGTGGTGTCTTCTGAGGAGCCAGAAATCAGCGAG 1191

QY 1141 GAGCTCAAGGACCTGATCTCTGAAGATGTTAGACAAGATCCCGAGACGAGAAATGGGGTG 1200
DB 1192 GAGCTCAAGGACCTGATCTCTGAAGATGTTAGACAAGATCCCGAGACGAGAAATGGGGTG 1251

QY 1201 CCAGACATCAAGTTGCACCCCTTGGGTGACCAAGACGGGGAGGAGCCCTTCCTTCGGAG 1260
DB 1252 CCAGACATCAAGTTGCACCCCTTGGGTGACCAAGACGGGGAGGAGCCCTTCCTTCGGAG 1311

QY 1261 GAGGAGCACTGACGCTGCTGAGGTGACAGAGGGGGAGGTAAAGAACTCAGTCAGGCTC 1320
DB 1312 GAGGAGCACTGACGCTGCTGAGGTGACAGAGGGGGAGGTAAAGAACTCAGTCAGGCTC 1371

QY 1321 ATCCCGAGCTGACCAACGCTGATCTCTGTAAGTGCATCTCAGGAAGCTTCCTTTGGG 1380
DB 1372 ATCCCGAGCTGACCAACGCTGATCTCTGTAAGTGCATCTCAGGAAGCTTCCTTTGGG 1431

QY 1381 AACCCGTTTGAGCCCGACGAGGAGGAGGAGCCATTCATCTCTCAGGAACCTA 1440
DB 1432 AACCCGTTTGAGCCCGACGAGGAGGAGGAGCCATTCATCTCTCAGGAACCTA 1491

QY 1441 CTGGTGAAGAAAGGGTTTGGTGAAGGGGCAAGAGCCAGAGCTCCCGCGCGTCCAGGAA 1500
DB 1492 CTGGTGAAGAAAGGGTTTGGTGAAGGGGCAAGAGCCAGAGCTCCCGCGCGTCCAGGAA 1551

QY 1501 GACGAGGCTGATCC 1515
DB 1552 GACGAGGCTGATCC 1566

RESULT 6

US-09-922-138-9
; Sequence 9, Application US/09922138
; Patent No. US20020061574A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Meyers, Rachel
; APPLICANT: Silos-Santiago, Inmaculada
; TITLE OF INVENTION: 16658, 14223, AND 16002, NOVEL HUMAN
; TITLE OF INVENTION: KINASES AND USES THEREFOR
; FILE REFERENCE: 38155-20030.00
; CURRENT APPLICATION NUMBER: US/09/922,138
; CURRENT FILING DATE: 2001-11-28
; PRIOR APPLICATION NUMBER: US 60/229,299
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 1683
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-922-138-9

Query Match 98.4%; Score 1490.6; DB 9; Length 1683;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1496; Conservative 0; Mismatches 9; Indels 0; Gaps 0;


```
; SEQ ID NO 43
; LENGTH: 1686
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1686)
US-10-391-364-43

Query Match      98.4%; Score 1490.6; DB 17; Length 1686;
Best local similarity 99.4%; Pred. No. 0;
Matches 1496; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 ATGGAGGGGGTCCAGCTGCTGCTGCCAGGATCCTCGGCAGAGCTGGTAGAACGGTGC 60
Db 1 ATGGAGGGGGTCCAGCTGCTGCTGCCAGGATCCTCGGCAGAGCTGGTAGAACGGTGC 60

QY 61 GCAGCCATCGATGTGACTCACTTGGAGGAGGACAGATGGTGGCCCCAGAGCCCTACTAGAAAC 120
Db 61 GCAGCCATCGATGTGACTCACTTGGAGGAGGACAGATGGTGGCCCCAGAGCCCTACTAGAAAC 120

QY 121 SGTGTGGACCCCCACACAGGGCCAGAGCTGCTCTGTGATCCCTGGCAGTACTTCAAGA 180
Db 121 SGTGTGGACCCCCACACAGGGCCAGAGCTGCTCTGTGATCCCTGGCAGTACTTCAAGA 180

QY 181 CTGCTCCAGCCGGGCTAGCCTCTCAGCCAGGAAGCTTTCCTACAGGAGCGGCCAGCA 240
Db 181 CTGCTCCAGCCGGGCTAGCCTCTCAGCCAGGAAGCTTTCCTACAGGAGCGGCCAGCA 240

QY 241 GGAAGCTATCTGAGGCGCAGGCTGGGGCTTTATGCCACGGGGCTCCAGCCACATCTCC 300
Db 241 GGAAGCTATCTGAGGCGCAGGCTGGGGCTTTATGCCACGGGGCTCCAGCCACATCTCC 300

QY 301 CCCCGGGCTGCGGAGGCCACCATCGAGTCCACACAGTGGCCATCTCAGATGCAGAG 360
Db 301 CCCCGGGCTGCGGAGGCCACCATCGAGTCCACACAGTGGCCATCTCAGATGCAGAG 360

QY 361 GACTGGCTGAGCTGAACAGTACAACTGACAGTGCAGATGGGCAAGGCTGCTACGCT 420
Db 361 GACTGGCTGAGCTGAACAGTACAACTGACAGTGCAGATGGGCAAGGCTGCTACGCT 420

QY 421 GTGGTGAAGCTGCTGACCAAGAAAGTGAAGACAGACATATGCAATGAAGTCTTTCC 480
Db 421 GTGGTGAAGCTGCTGACCAAGAAAGTGAAGACAGACATATGCAATGAAGTCTTTCC 480

QY 481 AAAAAGAGTTACTGAAGCAGTATGGCTTTCCAGCTCGCCCTCCCGGAGGGTCCCG 540
Db 481 AAAAAGAGTTACTGAAGCAGTATGGCTTTCCAGCTCGCCCTCCCGGAGGGTCCCG 540

QY 541 GCTGCCAGGAGGACCAAGCAGAGCTGCTGCCCTGGAGCGGGTGTACCAGGAGATT 600
Db 541 GCTGCCAGGAGGACCAAGCAGAGCTGCTGCCCTGGAGCGGGTGTACCAGGAGATT 600

QY 601 GCATCTGAAGAAGCTGACACAGTGAATGTGGTCAAACTGATCGAGTCTCTGGATGAC 660
Db 601 GCATCTGAAGAAGCTGACACAGTGAATGTGGTCAAACTGATCGAGTCTCTGGATGAC 660

QY 661 CCAGCTGAGGACAACTCTATTGTTGGTTGACCTCTCAGAAAGGGCCCTCATGGAA 720
Db 661 CCAGCTGAGGACAACTCTATTGTTGGTTGACCTCTCAGAAAGGGCCCTCATGGAA 720

QY 721 GTGCCCTGTGAAGCCCTTCTCGAGGAGCAAGCTCGCCTCTACTCGGGACGTCATC 780
Db 721 GTGCCCTGTGAAGCCCTTCTCGAGGAGCAAGCTCGCCTCTACTCGGGACGTCATC 780

QY 781 CTGGGCTCGAGTACTGCTGACGTCAGAGATCGTCCACAGGACATCAAGCATCCAC 840
Db 781 CTGGGCTCGAGTACTGCTGACGTCAGAGATCGTCCACAGGACATCAAGCATCCAC 840

QY 841 CTGCTCCTGGGGATGATGGCAGCTGAAGATCGCCGACTTTGGCGTCAGCAACAGTTT 900
Db 841 CTGCTCCTGGGGATGATGGCAGCTGAAGATCGCCGACTTTGGCGTCAGCAACAGTTT 900
```

```
QY 901 GAGGGAAACGACGCTCAGCTGTCCAGACGGGGGAAACCCAGCATTTATGGCCCCCGAG 960
Db 901 GAGGGAAACGACGCTCAGCTGTCCAGACGGGGGAAACCCAGCATTTATGGCCCCCGAG 960

QY 961 GCCATTTCTGATTCGGGCCAGAGCTTCAGTGGGAAGGCCCTTGGATGTATGGGCCACTGGC 1020
Db 961 GCCATTTCTGATTCGGGCCAGAGCTTCAGTGGGAAGGCCCTTGGATGTATGGGCCACTGGC 1020

QY 1021 GTACGCTTGTACTGCTTTTGTCTATGGGAAGTGGCCATTCATCGACGATTTATCTTGGCC 1080
Db 1021 GTACGCTTGTACTGCTTTTGTCTATGGGAAGTGGCCCTTCATCGACGATTTATCTTGGCC 1080

QY 1081 CTCCACAGGAAGATCAAGATGAGCCCGTGGTGTTCCTCGAGGAGCCAGAAATCAGCGAG 1140
Db 1081 CTCCACAGGAAGATCAAGATGAGCCCGTGGTGTTCCTCGAGGAGCCAGAAATCAGCGAG 1140

QY 1141 GAGCTCAAGGACCTGATCCTGAAGATGTTAGACAAGATCCCGAGACGAGAATTGGGGTG 1200
Db 1141 GAGCTCAAGGACCTGATCCTGAAGATGTTAGACAAGATCCCGAGACGAGAATTGGGGTG 1200

QY 1201 CCAGACATCAAGTTGACCCCTTGGGTGACCAAGAACGGGGAGGAGCCCTTCTTCCGGAG 1260
Db 1201 CCAGACATCAAGTTGACCCCTTGGGTGACCAAGAACGGGGAGGAGCCCTTCTTCCGGAG 1260

QY 1261 GAGGACACTGACGCGTGGTGGAGGTGACAGAGGGGGAGGTTAAGAACTCAGTCAGGCTC 1320
Db 1261 GAGGACACTGACGCGTGGTGGAGGTGACAGAGGGGGAGGTTAAGAACTCAGTCAGGCTC 1320

QY 1321 ATCCCGAGCTGACCAACGCTGATCCTGTGTAAGTCCATCTGAGGAAGCGTTCCTTTGGG 1380
Db 1321 ATCCCGAGCTGACCAACGCTGATCCTGTGTAAGTCCATCTGAGGAAGCGTTCCTTTGGG 1380

QY 1381 AACCCGTTGAGCCCCCAGCAGCGAGGAGGAGCGATCCATGCTCTCGAGAAACCTTA 1440
Db 1381 AACCCGTTGAGCCCCCAGCAGCGAGGAGGAGCGATCCATGCTCTCGAGAAACCTTA 1440

QY 1441 CTGGTGAAGAAAGGGTTTGGTGAAGGGGGCAAGAGCCAGAGCTCCCGGCGTCCAGGAA 1500
Db 1441 CTGGTGAAGAAAGGGTTTGGTGAAGGGGGCAAGAGCCAGAGCTCCCGGCGTCCAGGCT 1500

QY 1501 GACGA 1505
Db 1501 TACCA 1505
```

RESULT 8

```
US-09-922-138-7
; Sequence 7, Application US/09922138
; Patent No. US20020061574A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Meyers, Rachel
; APPLICANT: Silos-Santiago, Inmaculada
; TITLE OF INVENTION: 16658, 14223, AND 16002, NOVEL HUMAN
; TITLE OF INVENTION: KINASES AND USES THEREFOR
; FILE REFERENCE: 38155-20030.00
; CURRENT APPLICATION NUMBER: US/09/922,138
; CURRENT FILING DATE: 2001-11-28
; PRIOR APPLICATION NUMBER: US 60/229,299
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 2711
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (198)...(1880)
; NAME/KEY: misc.feature
; LOCATION: (1)-(2711)
; OTHER INFORMATION: n = A,T,C or G
US-09-922-138-7
```

Query Match		98.4%;	Score 1490.6;	DB 9;	Length 2711;	
Best Local Similarity		99.4%;	Pred. No. 0;			
Matches 1496;		Conservative	0;	Mismatches	9;	Indels 0; Gaps 0;
QY	1	ATGAGGGGGTCCAGTGTCTGTGTCAGAGATCCTCGGCGAGAGCTGGTAGAACGGGTG	60			
Db	198	ATGAGGGGGTCCAGTGTCTGTGTCAGAGATCCTCGGCGAGAGCTGGTAGAACGGGTG	257			
QY	61	GCAGCCATCATGTGACTCACTTGGAGAGGCAGATGGTGGCCCGAGAGCTTACTAGAAC	120			
Db	258	GCAGCCATCATGTGACTCACTTGGAGAGGCAGATGGTGGCCCGAGAGCTTACTAGAAC	317			
QY	121	GGTGTGACCCCGCCACACCGGGCGAGAGCTGCTGTGTGATCCCTGGCAGTACTCAAGA	180			
Db	318	GGTGTGACCCCGCCACACCGGGCGAGAGCTGCTGTGTGATCCCTGGCAGTACTCAAGA	377			
QY	181	CTGCTCCAGCCGGCTAGCTCTCAGCCAGGAAGCTTTCCCTACAGGAGCGCCAGCA	240			
Db	378	CTGCTCCAGCCGGCTAGCTCTCAGCCAGGAAGCTTTCCCTACAGGAGCGCCAGCA	437			
QY	241	GGAGCTATCTGGAGGGCGAGGCTGGGCTTATGCCACGGGGCTTGCACCATCTCC	300			
Db	438	GGAGCTATCTGGAGGGCGAGGCTGGGCTTATGCCACGGGGCTTGCACCATCTCC	497			
QY	301	CCCCGGGCTGGCGAGGCGCCACCATCGAGTCCACCAACCGTGGCCATCTCAGATGCAGAG	360			
Db	498	CCCCGGGCTGGCGAGGCGCCACCATCGAGTCCACCAACCGTGGCCATCTCAGATGCAGAG	557			
QY	361	GACTGCTGAGTGAAACAGTACAAGTGCAGAGTGAGATGGCAAGGGTGCTACGGT	420			
Db	558	GACTGCTGAGTGAAACAGTACAAGTGCAGAGTGAGATGGCAAGGGTGCTACGGT	617			
QY	421	GTGCTGAGGCTGGCTACACGAAGTGAACAGACACTATGCAATGAAGTCTCTTCC	480			
Db	618	GTGCTGAGGCTGGCTACACGAAGTGAACAGACACTATGCAATGAAGTCTCTTCC	677			
QY	481	AAAAAGAGTTACTGAAGCAGTAGTGGCTTTCACGTCGCCCTCCCGAGAGGGTCCCGAG	540			
Db	678	AAAAAGAGTTACTGAAGCAGTAGTGGCTTTCACGTCGCCCTCCCGAGAGGGTCCCGAG	737			
QY	541	GCTGCCAGGAGAGACAGCAAGCAGCTGTGCTGCCCTGGAGCGGGTGTTACAGAGATT	600			
Db	738	GCTGCCAGGAGAGACAGCAAGCAGCTGTGCTGCCCTGGAGCGGGTGTTACAGAGATT	797			
QY	601	GCCATCTGAAGAGCTGGACCACTGAATGTGTCAAACCTGATCGAGGCTCTGGATGAC	660			
Db	798	GCCATCTGAAGAGCTGGACCACTGAATGTGTCAAACCTGATCGAGGCTCTGGATGAC	857			
QY	661	CCAGCTGAGGACACCTCTATTGTTGGTGTGTTGACCTCCTCGAAGAGGGGCCCGTCATGAA	720			
Db	858	CCAGCTGAGGACACCTCTATTGTTGGTGTGTTGACCTCCTCGAAGAGGGGCCCGTCATGAA	917			
QY	721	GTGGCTGTGACAGCCCTTCTCGAGAGAGCAAGCTGCGCTCTACCTGCGGAGCTCATC	780			
Db	918	GTGGCTGTGACAGCCCTTCTCGAGAGAGCAAGCTGCGCTCTACCTGCGGAGCTCATC	977			
QY	781	CTGGCCCTCGAGTACTTGTGCTGCAGAGATCGTCCACAGGGACATCAAGCCATCAAC	840			
Db	978	CTGGCCCTCGAGTACTTGTGCTGCAGAGATCGTCCACAGGGACATCAAGCCATCAAC	1037			
QY	841	CTGCTCTCGGGGATGATGGGACAGTGAAGATCCCGACTTTGGCGTCAGCAACAGTTT	900			
Db	1038	CTGCTCTCGGGGATGATGGGACAGTGAAGATCCCGACTTTGGCGTCAGCAACAGTTT	1097			
QY	901	GAGGGGAACGACGCTCAGCTGTCAGCAACCGCGGGAACCCAGCATTCATGCCCCCGAG	960			
Db	1098	GAGGGGAACGACGCTCAGCTGTCAGCAACCGCGGGAACCCAGCATTCATGCCCCCGAG	1157			
QY	961	GCCATTTCTGATTCGCGCCAGAGCTTCAGTGGGAGGCTTGGATGTATGGCCACTGGC	1020			
Db	1158	GCCATTTCTGATTCGCGCCAGAGCTTCAGTGGGAGGCTTGGATGTATGGCCACTGGC	1217			

RESULT 9

US-10-391-364-41
; Sequence 41, Application US/10391364
; Publication No. US20040121349A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Meyers, Rachel E.
; APPLICANT: Carroll, Joseph M.
; APPLICANT: Cook, William James
; APPLICANT: Kapeller-Libermann, Rosana
; APPLICANT: Welch, Nadine S.
; APPLICANT: Bandaru, Rajasekhar
; TITLE OF INVENTION: NOVEL 27877, 18080, 14081, 32140, 50352,
; TITLE OF INVENTION: 16658, 14223, 16002, 50566, 65552 AND 65577 MOLECULES AND
; TITLE OF INVENTION: US\$ THEREFOR
; FILE REFERENCE: MEI03-0190NMIM
; CURRENT APPLICATION NUMBER: US/10/391,364
; CURRENT FILING DATE: 2003-03-18
; PRIOR APPLICATION NUMBER: US 09/950,370
; PRIOR FILING DATE: 2001-09-10
; PRIOR APPLICATION NUMBER: US 60/231,084
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: US 10/294,039
; PRIOR FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 60/338,587
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 10/266,035
; PRIOR FILING DATE: 2002-10-07
; PRIOR APPLICATION NUMBER: US 60/328,198
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: US 09/717,926
; PRIOR FILING DATE: 2000-11-21
; PRIOR APPLICATION NUMBER: US 60/214,707
; PRIOR FILING DATE: 2000-06-27
; PRIOR APPLICATION NUMBER: US 10/268,036
; PRIOR FILING DATE: 2002-10-09

QY	1021	GTCCAGTTGTACTGCTTTGTCTATGGGAAGTGCCCATTCATCGACGATTTTCATCTGCCC	1080
Db	1218	GTCCAGTTGTACTGCTTTGTCTATGGGAAGTGCCCGTTTCATCGACGATTTTCATCTGCCC	1277
QY	1081	CTCCACAGGAAGATCAAGATGAGCCCGTGGTGTTCCTGAGGAGCCAGAAATCAGCGAG	1140
Db	1278	CTCCACAGGAAGATCAAGATGAGCCCGTGGTGTTCCTGAGGAGCCAGAAATCAGCGAG	1337
QY	1141	GAGCTCAAGACCTGATCCCTGAAGATGTTAGACAAAGATCCCGAGACGAGAAATTTGGGGTG	1200
Db	1338	GAGCTCAAGACCTGATCCCTGAAGATGTTAGACAAAGATCCCGAGACGAGAAATTTGGGGTG	1397
QY	1201	CCAGACATCAAGTTGACCCCTTGGGTGACCAAGAACGGGAGGAGCCCTTCTTTCGGAG	1260
Db	1398	CCAGACATCAAGTTGACCCCTTGGGTGACCAAGAACGGGAGGAGCCCTTCTTTCGGAG	1457
QY	1261	GAGGAGCACTGCAGCGTGGTGGAGTGACAGAGGGGAGGTTAAGAACTCAGTTCAGGCTC	1320
Db	1458	GAGGAGCACTGCAGCGTGGTGGAGTGACAGAGGGGAGGTTAAGAACTCAGTTCAGGCTC	1517
QY	1321	ATCCCCAGCTGGACCAACCGTGTCTGCTGAAGTCCATGCTGAGGAAGCGTTCCCTTTGGG	1380
Db	1518	ATCCCCAGCTGGACCAACCGTGTCTGCTGAAGTCCATGCTGAGGAAGCGTTCCCTTTGGG	1577
QY	1381	AACCCGTTTCAGCCCCCAGGCAACGAGGAGGAAAGAGCATCCATGCTGCTCCAGGAAACCTA	1440
Db	1578	AACCCGTTTCAGCCCCCAGGCAACGAGGAGGAAAGAGCATCCATGCTGCTCCAGGAAACCTA	1637
QY	1441	CTGCTGAAGAAAGGTTTGGTGAAGGGGCAAGAGCCAGAGCTCCCGGGGTCAGAA	1500
Db	1638	CTGCTGAAGAAAGGTTTGGTGAAGGGGCAAGAGCCAGAGCTCCCGGGGTCAGAGCT	1697
QY	1501	GACGA 1505	
Db	1698	TACCA 1702	

```
; PRIOR APPLICATION NUMBER: US 60/327,820
; PRIOR FILING DATE: 2001-10-09
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 93
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 41
; LENGTH: 2711
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (198)...(1883)
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)...(2711)
; OTHER INFORMATION: n = A,T,C or G
US-10-391-364-41

Query Match          98.4%; Score 1490.6; DB 17; Length 2711;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1496; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY      1  ATGAGGGGGTCCAGCTGTCTCTGCGAGGATCTCTCGGGCAGAGCTGTGTAGAACGGGTG 60
DB      198  ATGGAGGGGGTCCAGCTGTCTCTGCGAGGATCTCTCGGGCAGAGCTGTGTAGAACGGGTG 257

QY      61  GCAGCCATCGATGTGACTCACTTGGAGGAGGACAGATGTTGGCCACAGAGCTACTAGAAAC 120
DB      258  GCAGCCATCGATGTGACTCACTTGGAGGAGGACAGATGTTGGCCACAGAGCTACTAGAAAC 317

QY      121  GGTGTGGAACCCGCCACACGGGCCAGAGTGCCTCTGTGATCCCTGGCAGTACTTCAAGA 180
DB      318  GGTGTGGAACCCGCCACACGGGCCAGAGTGCCTCTGTGATCCCTGGCAGTACTTCAAGA 377

QY      181  CTGCTCCAGCCGGCTAGCCTCTCAGCCAGGAAGCTTTCCTACAGAGCGGCCAGCA 240
DB      378  CTGCTCCAGCCGGCTAGCCTCTCAGCCAGGAAGCTTTCCTACAGAGCGGCCAGCA 437

QY      241  GGAAGCTATCTGGAGGCGCAGGCTGGGCCCTTATGCCACGGGGCTGCCAGCCACATCTCC 300
DB      438  GGAAGCTATCTGGAGGCGCAGGCTGGGCCCTTATGCCACGGGGCTGCCAGCCACATCTCC 497

QY      301  CCCCAGGCTGGCGGAGGCCACCATCGAGTCCACACCGTGCCCATCTCAGATCGAG 360
DB      498  CCCCAGGCTGGCGGAGGCCACCATCGAGTCCACACCGTGCCCATCTCAGATCGAG 557

QY      361  GACTCGTGCAGTGAACAGTACAGCTGCGAGTGAGATTGGCAGGCTGCTACGTT 420
DB      558  GACTCGTGCAGTGAACAGTACAGCTGCGAGTGAGATTGGCAGGCTGCTACGTT 617

QY      421  GTGGTGAGGCTGGCCCTACACGAAAGTGAAGACAGACACTATGCAATGAAAGTCTCTTC 480
DB      618  GTGGTGAGGCTGGCCCTACACGAAAGTGAAGACAGACACTATGCAATGAAAGTCTCTTC 677

QY      481  AAAAAGAGTTACTGAAGCAGTATGGCTTTTCCAGTGGCCCTCCCGGAGAGGGTCCAG 540
DB      678  AAAAAGAGTTACTGAAGCAGTATGGCTTTTCCAGTGGCCCTCCCGGAGAGGGTCCAG 737

QY      541  GCTGCCAGGGAGGACCCAGCAGAGCTGCTGCCCTGGAGGGGTGTACACGAGATT 600
DB      738  GCTGCCAGGGAGGACCCAGCAGAGCTGCTGCCCTGGAGGGGTGTACACGAGATT 797

QY      601  GCCATCTTGAAGAAGCTGACACCGTGAATGTGGTCAAACTGATCGAGGTCTCGATGAC 660
DB      798  GCCATCTTGAAGAAGCTGACACCGTGAATGTGGTCAAACTGATCGAGGTCTCGATGAC 857

QY      661  CCAGCTGAGGACAACTTATTGTGTTTGTACCTCTGAGAAAGGGGCCGTGATGNA 720
DB      858  CCAGCTGAGGACAACTTATTGTGTTTGTACCTCTGAGAAAGGGGCCGTGATGNA 917

QY      721  GTGGCTGTGACAAAGCCCTTCTCGGAGGAGCAAGCTCGCCTCTACTCGGGACGTCATC 780
DB      918  GTGGCTGTGACAAAGCCCTTCTCGGAGGAGCAAGCTCGCCTCTACTCGGGACGTCATC 977

QY      781  CTGGGGCTCGACTACTTGCACCTGCCAGAAAGATCGTCCACAGGACATCAAGCCATCCAAC 840
DB      978  CTGGGGCTCGACTACTTGCACCTGCCAGAAAGATCGTCCACAGGACATCAAGCCATCCAAC 1037

QY      841  CTGCTCCTGGGGATGATGGGCACAGTGAAGATCGCCGACTTTTGGCGTCAGCAACCACTTT 900
DB      1038  CTGCTCCTGGGGATGATGGGCACAGTGAAGATCGCCGACTTTTGGCGTCAGCAACCACTTT 1097

QY      901  GAGGGGAACGACGCTCAGCTGTCCAGCACGGCGGGAACCCACGATTCATGGCCCCCGAG 960
DB      1098  GAGGGGAACGACGCTCAGCTGTCCAGCACGGCGGGAACCCACGATTCATGGCCCCCGAG 1157

QY      961  GCATTTCTGATTCCGGCCAGAGCTTCAGTGGGAAGCCCTTGATGTATGGGCCACTGGC 1020
DB      1158  GCATTTCTGATTCCGGCCAGAGCTTCAGTGGGAAGCCCTTGATGTATGGGCCACTGGC 1217

QY      1021  GTACGTTTGTACTGCTTTCTTATGGGAAGTGCCTTTCATCGACGATTCATCTCTGGCC 1080
DB      1218  GTACGTTTGTACTGCTTTCTTATGGGAAGTGCCTTTCATCGACGATTCATCTCTGGCC 1277

QY      1081  CTCCACAGGAAGATCAAGATGAGCCCGTGGTGTTCCTGAGGAGCCAGAAATCAGCGAG 1140
DB      1278  CTCCACAGGAAGATCAAGATGAGCCCGTGGTGTTCCTGAGGAGCCAGAAATCAGCGAG 1337

QY      1141  GAGCTCAAGGACCTGATCCTGAAGATGTTAGACAAGAATCCCGAGACGAGAAATTGGGGTG 1200
DB      1338  GAGCTCAAGGACCTGATCCTGAAGATGTTAGACAAGAATCCCGAGACGAGAAATTGGGGTG 1397

QY      1201  CCAGACATCAAGTTGACCCCTTTGGGTGACCAAGAACGGGGAGAGCCCTTCTCTCGGAG 1260
DB      1398  CCAGACATCAAGTTGACCCCTTTGGGTGACCAAGAACGGGGAGAGCCCTTCTCTCGGAG 1457

QY      1261  GAGGAGCTCGCAGCGTGGTGGAGGTGACAGAGGGGGAGGTTAAGAACTCAGTCAGGCTC 1320
DB      1458  GAGGAGCTCGCAGCGTGGTGGAGGTGACAGAGGGGGAGGTTAAGAACTCAGTCAGGCTC 1517

QY      1321  ATCCCCAGCTGGACCAACGGTGATCTGGTGAAGTCCATCTCAGGAAGCGTTCCTTTGGG 1380
DB      1518  ATCCCCAGCTGGACCAACGGTGATCTGGTGAAGTCCATCTCAGGAAGCGTTCCTTTGGG 1577

QY      1381  AACCCGTTTACCCCCAGCACCGAGGGGAAGCGCATCATGTCTCCAGAAACCTA 1440
DB      1578  AACCCGTTTACCCCCAGCACCGAGGGGAAGCGCATCATGTCTCCAGAAACCTA 1637

QY      1441  CTGGTGAAGAGGGTTTGGTGAAGGGGCGAGAGCCAGAGTCCCGGCGTCCAGAA 1500
DB      1638  CTGGTGAAGAGGGTTTGGTGAAGGGGCGAGAGTCCCGGCGTCCAGAA 1697

QY      1501  GACGA 1505
DB      1698  TACCA 1702

RESULT 10
US-10-168-582-14
; Sequence 14, Application US/10168582
; Publication No. US20040058426A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: YANG, Junming
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: BURFORD, Neil
; APPLICANT: AU-YOUNG, Janice
; APPLICANT: LU, Dyung Aina M.
; APPLICANT: REDDY, Roopa
; APPLICANT: YUE, Henry
; APPLICANT: YAO, Monique G.
; APPLICANT: LAL, Preeti
; APPLICANT: KAHN, Farrah A.
; TITLE OF INVENTION: HUMAN KINASES
; FILE REFERENCE: PI-0002 PCT
; CURRENT APPLICATION NUMBER: US/10/168,582
```

```
; CURRENT FILING DATE: 2002-06-20
; PRIOR APPLICATION NUMBER: 60/172,066; 60/176,107; 60/176,107; 60/177,731
; PRIOR FILING DATE: 1999-12-23; 2000-01-14; 2000-01-14; 2000-01-21
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PERL Program
; SEQ ID NO 14
; LENGTH: 3501
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20040058426A1 2041716CB1
; US-10-168-582-14

Query Match      96.2%; Score 1457.2; DB 13; Length 3501;
Best Local Similarity 97.9%; Pred. No. 0;
Matches 1509; Conservative 0; Mismatches 3; Indels 30; Gaps 2;

QY 1 ATGAGGGGGTCCAGCTGTCTGTGCGCAGGATCCTCGGGCAGAGCTGTGTAGAACGGGTG 60
Db 1 ATGAGGGGGTCCAGCTGTCTGTGCGCAGGATCCTCGGGCAGAGCTGTGTAGAACGGGTG 124

QY 61 GCACCCATCGATGTGACTCACTTGGAGAGGCGAGATGGTGGCCCGAGAGCTTACTAGAAC 120
Db 61 GCACCCATCGATGTGACTCACTTGGAGAGGCGAGATGGTGGCCCGAGAGCTTACTAGAAC 184

QY 121 GGTGTGACCCGCCACACACCGGGCCAGAGCTGCCTCTGTGTATCCCTGGCAGTACTTCAAGA 180
Db 121 GGTGTGACCCGCCACACACCGGGCCAGAGCTGCCTCTGTGTATCCCTGGCAGTACTTCAAGA 244

QY 181 CTGCTCCAGCCGGCTAGCCTCTCAGCCAGAGAACTTTCCTTACAGGAGCGGCCAGCA 240
Db 181 CTGCTCCAGCCGGCTAGCCTCTCAGCCAGAGAACTTTCCTTACAGGAGCGGCCAGCA 304

QY 241 GGAAGCTATCTGGAGGGCGAGGCTGGGCTTATGCCAGGGGCTGCCAGCCATCTCC 300
Db 241 GGAAGCTATCTGGAGGGCGAGGCTGGGCTTATGCCAGGGGCTGCCAGCCATCTCC 364

QY 301 CCCGGGCTTGGAGGGCGAGGCTGCCAGTCCAGTCCACACCGAGTGGCCATCTCAGATGCGAG 360
Db 301 CCCGGGCTTGGAGGGCGAGGCTGCCAGTCCAGTCCACACCGAGTGGCCATCTCAGATGCGAG 424

QY 361 GACTGCGTCAAGTGAACAGTCAAGCTGAGAGTGAAGTGGCA----- 406
Db 361 GACTGCGTCAAGTGAACAGTCAAGCTGAGAGTGAAGTGGCA----- 484

QY 407 -----AGGCTGCTACGGTGTGGTGGAGGCTGCCAGTGGAGTGGAGTGGAGTGGAG 453
Db 407 -----AGGCTGCTACGGTGTGGTGGAGGCTGCCAGTGGAGTGGAGTGGAGTGGAG 544

QY 454 AGACACTATGCAATGAAAGTCCCTTCCAAAAGAAAGTTACTGAAAGCAGTATGGCTTTCCA 513
Db 454 AGACACTATGCAATGAAAGTCCCTTCCAAAAGAAAGTTACTGAAAGCAGTATGGCTTTCCA 604

QY 514 CGTCGCCCTCCCGAGAGGGTCCAGGCTGCCAGGGAGGACAGCAGCAGAGCTGCTG 573
Db 514 CGTCGCCCTCCCGAGAGGGTCCAGGCTGCCAGGGAGGACAGCAGCAGAGCTGCTG 664

QY 574 CCCTGGAGCGGGTGTACAGGAGATGGCCATCTGAGAGCTGGAGCCAGCTGAATGTG 633
Db 574 CCCTGGAGCGGGTGTACAGGAGATGGCCATCTGAGAGCTGGAGCCAGCTGAATGTG 724

QY 634 GTCAAACTGATCGAGGTCCTGGATGAGCCAGCTGAGGACCACTCTATTTGGTGTGTGAC 693
Db 634 GTCAAACTGATCGAGGTCCTGGATGAGCCAGCTGAGGACCACTCTATTTGGTGTGTGAC 781

QY 694 CTCCTGAGAAAGGGGCCCGCTCATGGAAGTGCCTGTGACAAAGCCCTTCTCGGAGGACAA 753
Db 694 CTCCTGAGAAAGGGGCCCGCTCATGGAAGTGCCTGTGACAAAGCCCTTCTCGGAGGACAA 841

QY 754 GCTCGCCTCTACCTGCGGGAGCGTCACTCGGGGCTCGAGTACTTGCACGAGAGATC 813
Db 754 GCTCGCCTCTACCTGCGGGAGCGTCACTCGGGGCTCGAGTACTTGCACGAGAGATC 901

QY 842 GCTCGCCTCTACCTGCGGGAGCGTCACTCGGGGCTCGAGTACTTGCACGAGAGATC 901
```

```
RESULT 11
US-10-092-900A-53
; Sequence 53, Application US/10092900A
; Publication No. US20040043382A1
; GENERAL INFORMATION:
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Spytek, Kimberlly A.
; APPLICANT: Shenoy, Suresh G.
; APPLICANT: Taupier Jr., Raymond J.
; APPLICANT: Pena, Carol E.A.
; APPLICANT: Li, Li
; APPLICANT: Zerhusen, Bryan D.
; APPLICANT: Gusev, Vladimir Y.
; APPLICANT: Ji, Weizhen
; APPLICANT: Gorman, Linda
; APPLICANT: Miller, Charles E.
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Patturajan, Meera
; APPLICANT: Gangolli, Esha A.
; APPLICANT: Vernet, Corine A.M.
; APPLICANT: Guo, Xiaojia Sasha
; APPLICANT: Tchernev, Velizar T.
; APPLICANT: Fernandes, Elma R.
```

```
QY 814 GTCCACAGGACATCAAGCCATCAACCTGCTCTCTGGGGATGATGGCAGTGAAGATC 873
Db 814 GTCCACAGGACATCAAGCCATCAACCTGCTCTCTGGGGATGATGGCAGTGAAGATC 961

QY 874 GCCGACTTTGGCGTCAAGAACCAAGTTTGGGGGAAACGACGCTCAGCTGTCTCCAGCAGCGG 933
Db 874 GCCGACTTTGGCGTCAAGAACCAAGTTTGGGGGAAACGACGCTCAGCTGTCTCCAGCAGCGG 1021

QY 934 GGAACCCAGCATTCATGGCCCCCGAGCCATTTCTGATTTCCGSCCAGAGCTTCAGTGGG 993
Db 934 GGAACCCAGCATTCATGGCCCCCGAGCCATTTCTGATTTCCGSCCAGAGCTTCAGTGGG 1081

QY 994 AAGSCCTTGTATGTATGGGCCATCTGGCGTCAAGCTTGTACTGCTTGTCTATGGGAAGTGC 1053
Db 994 AAGSCCTTGTATGTATGGGCCATCTGGCGTCAAGCTTGTACTGCTTGTCTATGGGAAGTGC 1141

QY 1054 CCATTCAATCAAGATTTTCATCTTGGCCCTCCACAGGAAGATCAAGAAATGAGCCCGTGGT 1113
Db 1054 CCATTCAATCAAGATTTTCATCTTGGCCCTCCACAGGAAGATCAAGAAATGAGCCCGTGGT 1201

QY 1114 TTTCTCAGGAGCAGAAATCAGCGAGAGCTCAAGAACCTGATCCCTGAAGATGTTAGAC 1173
Db 1114 TTTCTCAGGAGCAGAAATCAGCGAGAGCTCAAGAACCTGATCCCTGAAGATGTTAGAC 1261

QY 1174 AAGAATCCCGAGAGCAGAAATTTGGGTGCCAGACATCAAGTTGCAACCTTGGGTGACCAAG 1233
Db 1174 AAGAATCCCGAGAGCAGAAATTTGGGTGCCAGACATCAAGTTGCAACCTTGGGTGACCAAG 1321

QY 1234 AAGCGGAGGAGGCCCTTCTTCCGAGGAGAGCAGCTCAGCGGTGGTGGAGGTGACAGAG 1293
Db 1234 AAGCGGAGGAGGCCCTTCTTCCGAGGAGAGCAGCTCAGCGGTGGTGGAGGTGACAGAG 1381

QY 1294 GGGGAGGTTAAGAACTCAGTCAGGCTCATCCCGAGCTGGACCGGTGATCTCTGTTGAAG 1353
Db 1294 GGGGAGGTTAAGAACTCAGTCAGGCTCATCCCGAGCTGGACCGGTGATCTCTGTTGAAG 1441

QY 1354 TCCATGCTGAGGAGCGTCTCTTTGGGAAACCCGTTTGGAGCCAGGACCGAGGAGGAGAG 1413
Db 1354 TCCATGCTGAGGAGCGTCTCTTTGGGAAACCCGTTTGGAGCCAGGACCGAGGAGGAGAG 1501

QY 1414 CGATCCATGTCTCTCAGGAAACCTTACTCTGTGAAAGAGGGTTTGTGAGGGGGCAAG 1473
Db 1414 CGATCCATGTCTCTCAGGAAACCTTACTCTGTGAAAGAGGGTTTGTGAGGGGGCAAG 1561

QY 1474 AGCCGAGCTCCCGCGCTCCAGGAGAGCAGAGCTGCATCC 1515
Db 1474 AGCCGAGCTCCCGCGCTCCAGGAGAGCAGAGCTGCATCC 1603
```



```
; APPLICANT: Casman, Stacie J.
; APPLICANT: Malyankar, Uriel M.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Liu, Yi
; APPLICANT: Anderson, David W.
; APPLICANT: Spaderina, Steven K.
; APPLICANT: Catterton, Elina
; APPLICANT: Leite, Mario W.
; APPLICANT: Zhong, Haihong
; APPLICANT: Alsobrook, John P.
; APPLICANT: Lopley, Denise M.
; APPLICANT: Rieger, Daniel K.
; APPLICANT: Burgess, Catherine E.
; TITLE OF INVENTION: No. US20040043382A1el Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-290C
; CURRENT APPLICATION NUMBER: US/10/092,900A
; CURRENT FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: USSN 60/274,322
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: USSN 60/283,675
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: USSN 60/338,092
; PRIOR FILING DATE: 2001-12-03
; PRIOR APPLICATION NUMBER: USSN 60/274,281
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: USSN 60/274,191
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: USSN 60/325,681
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: USSN 60/304,354
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: USSN 60/279,995
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: USSN 60/294,899
; PRIOR FILING DATE: 2001-05-31
; PRIOR APPLICATION NUMBER: USSN 60/287,424
; PRIOR FILING DATE: 2001-04-30
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 768
; SEQ ID NO 53
; LENGTH: 1547
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (20)..(1529)
; US-10-092-900A-53

Query Match          96.1%; Score 1455.8; DB 13; Length 1547;
Best Local Similarity 98.5%; Pred. No. 0;
Matches 1492; Conservative 0; Mismatches 17; Indels 6; Gaps 2;

QY      1  ATGGAGGGGGTCCAGCTCTGTGTGCCAGGATCTCTCGGGCAGAGCTGTGTAGAACGGGTG 60
Db      1  |||||
QY      20  ATGGAGGGGGTCCAGCTCTGTGTGCCAGGATCTCTCGGGCAGAGCTGTGTAGAACGGGTG 79
Db      20  |||||
QY      61  GCAGCCATCATGTGACTCACTTGGAGGAGGAGATGGTGGCCACAGAGCTTACTAGAAAC 120
Db      61  |||||
QY      80  GCAGCCATCATGTGACTCACTTGGAGGAGGAGATGGTGGCCACAGAGCTTACTAGAAAC 139
Db      80  |||||
QY      121  GGTGTGGACCCCCACACGGGCCAGAGCTGTGTATCCCTGGCAGTACTTCAAGA 180
Db      121  |||||
QY      140  GGTGTGGACCCCCACACGGGCCAGAGCTGTGTATCCCTGGCAGTACTTCAAGA 199
Db      140  |||||
QY      181  CTGCTCCAGCGGGCTAGCCTCTCAGCAGAGCTTCCCTACAGAGCGGCCAGCA 240
Db      181  |||||
QY      200  CTGCTCCAGCGGGCTAGCCTCTCAGCAGAGCTTCCCTACAGAGCGGCCAGCA 259
Db      200  |||||
QY      241  GGAAGCTATCTGAGGCGCAGGCTGGGCTTATGCCAGCGGGCTGCCAGCCACATCTCC 300
Db      241  |||||
QY      260  GGAAGCTATCTGAGGCGCAGGCTGGGCTTATGCCAGCGGGCTGCCAGCCACATCTCC 319
Db      260  |||||
QY      301  CCCCAGGCTGGGGAGGGCCACCATCGAGTCCCACCACGTGGCCATCTCAGATGCAGAG 360
Db      301  |||||

320  CCCCAGGCTGGGGAGGGCCACCATCGAGTCCCACCACGTGGCCATCTCAGATGCAGAG 379
361  GACTCGTGCAGCTGAACAGGTACAAGCTGCAGAGTGAGATTGGCAAGGGTGCTTACGGT 420
380  GACTCGTGCAGCTGAACAGGTACAAGCTGCAGAGTGAGATTGGCAAGGGTGCTTACGGT 439
421  GTGGTGAGGCTGGCCCTACAAAGAGTGNAGACAGACACTATGCATGAAGTCCCTTCC 480
440  GTGGTGAGGCTGGCCCTACAAAGAGTGAAGACAGACACTATGCATGAAGTCCCTTCC 499
481  AAAAAGAGTGTACTGAAGCAGTATGGCTTTTCCACGTGCGCCTCCCCCGAGAGGGTCCCAG 540
500  AAAAAGAGTGTACTGAAGCAGTATGGCTTTTCCACGTGCGCCTCCCCCGAGAGGGTCCCAG 559
541  GTGCCCGAGGGAGGACCGACAGAGCTGTGCGCCCTGGAGCGGGTGTACAGAGATT 600
560  GCTGCCCGAGGGAGGACCGACAGAGCTGTGCGCCCTGGAGCGGGTGTACAGAGATT 619
601  GCATCTCTGAAGAGCTGGACCAAGTGAATGTGTTCAAACTGATCGAGGTACTGGATGAC 660
620  GCATCTCTGAAGAGCTGGACCAAGTGAATGTGTTCAAACTGATCGAGGTACTGGATGAC 679
661  CCAGCTGAGGACAACTCTTATTTGGTGTGTGACCTCTCGAGAAAGGGGGCCCTCATGGAA 720
680  CCAGCTGAGGACAACTCTTATTTGGTGTGTGACCTCTCGAGAAAGGGGGCCCTCATGGAA 736
721  GTGCCCTGTGACAAGCCCTTTCGGAGAGGAAAGCTGCGCTCTACTCTGCGGAGCTGATC 780
737  GTGCCCTGTGACAAGCCCTTTCGGAGAGGAAAGCTGCGCTCTACTCTGCGGAGCTGATC 796
781  CTGGGCTCTGAGTCTGACAGTGCAGAGAGATGCTGCCAGAGGACATCAAGCATCAAC 840
797  CTGGGCTCTGAGTCTGACAGTGCAGAGAGATGCTGCCAGAGGACATCAAGCATCAAC 856
841  CTGCTCTCTGGGGATGATGGGCAAGTGAAGATCGCGGACTTTGGGCTCAGCAACCAAGTTT 900
857  CTGCTCTCTGGGGATGATGGGCAAGTGAAGATCGCGGACTTTGGGCTCAGCAACCAAGTTT 916
901  GAGGGAGACGAGCTCAGCTGTCAGACAGCGGGGGAAACCCAGCATTCATGCCCCCGAG 960
917  GAGGGAGACGAGCTCAGCTGTCAGACAGCGGGGGAAACCCAGCATTCATGCCCCCGAG 976
961  GCATTTCTGATTTCGGCCAGAGCTTCACTGGGAAGGCTTGGATGTATGGGCCACTGGC 1020
977  GCCATTTCTGATTTCGGCCAGAGCTTCACTGGGAAGGCTTGGATGTATGGGCCACTGGC 1033
1021  GTCAGCTGTACTGCTTTGTCTATGGGAAGTCCCATTCATCGACGATTCATTCCTGGCC 1080
1034  GTCAGCTGTACTGCTTTGTCTATGGGAAGTCCCATTCATCGACGATTCATTCCTGGCC 1093
1081  CTCACAGGAAGATCAAGATGAGCCCGTGGTGTTCCTGAGAGCCAGAAATCAGCGAG 1140
1094  CTCACAGGAAGATCAAGATGAGCCCGTGGTGTTCCTGAGAGCCAGAAATCAGCGAG 1153
1141  GAGCTCAAGGACCTGATCCTGAAGATGTAGACAAGATCCCGAGACGAGAATTTGGGGTG 1200
1154  GAGCTCAAGGACCTGATCCTGAAGATGTAGACAAGATCCCGAGACGAGAATTTGGGGTG 1213
1201  CCAGACATCAAGTTGGACCTTGGGTGACCAAGACCGGGAGGAGCCCTTCTCTCGGAG 1260
1214  CCAGACATCAAGTTGGACCTTGGGTGACCAAGACCGGGAGGAGCCCTTCTCTCGGAG 1273
1261  GAGGAGCATCTGAGCGTGTGGAGGTGACAGAGAGCGGGAGGAGCCCTTCTCTCGGAG 1320
1274  GAGGAGCATCTGAGCGTGTGGAGGTGACAGAGAGCGGGAGGAGTTAAGACTCAGTCAAGCTC 1333
1321  ATCCCCAGCTGACCAAGCTGATCTCTGGTGAAGTCCATCTGAGGAAGGCTTCTCTTGGG 1380
1334  ATCCCCAGCTGACCAAGCTGATCTCTGGTGAAGTCCATCTGAGGAAGGCTTCTCTTGGG 1393
1381  AACCCGTTGAGCCCCCAGGACGAGGAGGAGCGGATCCATCTGCTCCAGGAACCTTA 1440
```

Db 1394 AACCGTTTGAGCCCAAGCACGAGGAGGAGGAGGATCCATGTCTGTCTCCAGGAAACCTA 1453
Qy 1441 CTGTGTAAGAGAGGTTTGTGAAGGGGCAAGAGCCACAGAGCTCCCGGGCTCCAGGAA 1500
Db 1454 CTGTGTAAGAGAGGTTTGTGAAGGGGCAAGAGCCACAGAGCTCCCGGGCTCCAGGAA 1513
Qy 1501 GACGAGCTGCATCC 1515
Db 1514 GACGAGCTGCATCC 1528

RESULT 12

US-10-092-900A-57
; Sequence 57, Application US/10092900A
; Publication No. US20040043382A1
; GENERAL INFORMATION:
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Shenoy, Suresh G.
; APPLICANT: Taupier Jr., Raymond J.
; APPLICANT: Pena, Carol E.A.
; APPLICANT: Li, Li
; APPLICANT: Zerhusen, Bryan D.
; APPLICANT: Gusev, Vladimir Y.
; APPLICANT: Ji, Weizhen
; APPLICANT: Gorman, Linda
; APPLICANT: Miller, Charles E.
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Patturajan, Meera
; APPLICANT: Gangolli, Esha A.
; APPLICANT: Vernet, Corine A.M.
; APPLICANT: Guo, Xiaojia Sasha
; APPLICANT: Tchernev, Velizar T.
; APPLICANT: Fernandes, Elma R.
; APPLICANT: Casman, Stacie J.
; APPLICANT: Malyankar, Uriel M.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Liu, Yi
; APPLICANT: Anderson, David W.
; APPLICANT: Spaderna, Steven K.
; APPLICANT: Carterton, Elina
; APPLICANT: Leite, Mario W.
; APPLICANT: Zhong, Haihong
; APPLICANT: Alsobrook, John P.
; APPLICANT: Lepley, Denise M.
; APPLICANT: Rieget, Daniel K.
; APPLICANT: Burgess, Catherine E.
; TITLE OF INVENTION: No. US20040043382A1el Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-290C
; CURRENT APPLICATION NUMBER: US/10/092,900A
; CURRENT FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: USN 60/274,322
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: USN 60/283,675
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: USN 60/338,092
; PRIOR FILING DATE: 2001-12-03
; PRIOR APPLICATION NUMBER: USN 60/274,281
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: USN 60/274,191
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: USN 60/325,681
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: USN 60/304,354
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: USN 60/279,995
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: USN 60/294,899
; PRIOR FILING DATE: 2001-05-31
; PRIOR APPLICATION NUMBER: USN 60/287,424
; PRIOR FILING DATE: 2001-04-30
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 768

; SEQ ID NO 57
; LENGTH: 1725
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (52)..(1681)
US-10-092-900A-57

Query Match 91.2%; Score 1381.4; DB 13; Length 1725;
Best Local Similarity 92.6%; Pred. No. 0;
Matches 1509; Conservative 0; Mismatches 6; Indels 114; Gaps 1;

Qy 1 ATGGAGGGGGTCCAGCTGTCTGTGCGCAGGATCCTCGGGCAGAGCTGGTAGAACGGGTG 60
Db |||||
Qy 52 ATGGAGGGGGTCCAGCTGTCTGTGCGCAGGATCCTCGGGCAGAGCTGGTAGAACGGGTG 111
Db |||||
Qy 61 GCAGCCATCGATGTGACTCACTTGGAGGAGGAGATGGTGCCCGAGAGCCTTACTAGAAAC 120
Db |||||
Qy 112 GCAGCCATCGATGTGACTCACTTGGAGGAGGAGATGGTGCCCGAGAGCCTTACTAGAAAC 171
Db |||||
Qy 121 GGTGTGACCCGCCACCCAGCGGCCAGAGCTGCTCTGTGATCCCTGCGAGTACTTCAAGA 180
Db |||||
Qy 172 GGTGTGACCCGCCACCCAGCGGCCAGAGCTGCTCTGTGATCCCTGCGAGTACTTCAAGA 231
Db |||||
Qy 181 CTGCTCCAGCCGGGCTAGCCTCTCAGCCAGGAAGCTTTCCCTACAGGAGCGGCCAGCA 240
Db |||||
Qy 232 CTGCTCCAGCCGGGCTAGCCTCTCAGCCAGGAAGCTTTCCCTACAGGAGCGGCCAGCA 291
Db |||||
Qy 241 GGAAGCTATCTGAGGCGCAGGCTGGGCGCTTATGCGCAGGGGCTGCCAGCACATCTCC 300
Db |||||
Qy 292 GGAAGCTATCTGAGGCGCAGGCTGGGCGCTTATGCCACGGGGCTGCCAGCACATCTCC 351
Db |||||
Qy 301 CCCGGGCTTGGCGGAGGCCACCATCGAGTCCCACACGTGGCCATCTCAGANTGCAGAG 360
Db |||||
Qy 352 CCCGGGCTTGGCGGAGGCCACCATCGAGTCCCACACGTGGCCATCTCAGANTGCAGAG 411
Db |||||
Qy 361 GACTGCGTGAGCTGAAACCAGTACAGCTGCAGAGTGAGATTGGCAAGGGTGCTTACGGT 420
Db |||||
Qy 412 GACTGCGTGAGCTGAAACCAGTACAGCTGCAGAGTGAGATTGGCAAGGGTGCTTACGGT 471
Db |||||
Qy 421 GTGTGAGGCTGGCCTTACAGCAAGTGAAGACAGACACTATGCAATGAAAGTCTTTTCC 480
Db |||||
Qy 472 GTGTGAGGCTGGCCTTACAGCAAGTGAAGACAGACACTATGCAATGAAAGTCTTTTCC 531
Db |||||
Qy 481 AAAAAGAGTTACTGAGAGCATGTGGCTTTCCAGCTGCCCTCCCGAGAGGGTCCAG 540
Db |||||
Qy 532 AAAAAGAGTTACTGAGAGCATGTGGCTTTCCAGCTGCCCTCCCGAGAGGGTCCAG 591
Db |||||
Qy 541 GCTGCCCGAGGAGGAGCCAGCAAGCAGCTGCTGCCCTGGAGCGGGTGTACAGAGATT 600
Db |||||
Qy 592 GCTGCCCGAGGAGGAGCCAGCAAGCAGCTGCTGCCCTGGAGCGGGTGTACAGAGATT 651
Db |||||
Qy 601 GCATCTCTGAAGAGCTGAGCCAGTGAATGTGGTCAAACTGATCGAGGTCCTGGATGAC 660
Db |||||
Qy 652 GCCATCTCTGAAGAGCTGAGCCAGCTGAATGTGGTCAAACTGATCGAGGTCCTGGATGAC 711
Db |||||
Qy 661 CCAGCTGAGGACAACTCTTATTG----- 684
Db |||||
Qy 712 CCGGCTGAGGACAACTCTTATTGCGCCCTGCAGAACCGGCCCGAGANTATCCAGTTAGAT 771
Db |||||
Qy 685 ----- 684
Db |||||
Qy 772 TCAACAAATATGCGCAAGTCCCACTCCCTGCTTCCCTCTGAGCAGCAAGACAGTGGATCC 831
Db |||||
Qy 685 -----GTGTTTGACCTCTCAGAAAAGGGGCCCGTCTATGGAAGTCCCG 726
Db |||||
Qy 832 ACSTGGGCTCGGGCTCAGTGTGTTGACCTCTCAGAAAAGGGGCCCGTCTATGGAAGTCCCG 891
Db |||||
Qy 727 TGTGACAAGCCCTTCTCGGAGGAGCAGCTCGCTCTACCTCGGGGACGTCATCTCTGGC 786
Db |||||
Qy 892 TGTGACAAGCCCTTCTCGGAGGAGCAAGCTCGCTCTTACCTGCGGGACGTCATCTCTGGC 951
Db |||||

```
QY 787 CTCGACTACTTGCACTGCCAGAGATGCTCCACAGGACATCAAGCCATCCACCTGCTC 846
Db 952 CTCGAGTACTTGCACTGCCAGAGATGCTCCACAGGACATCAAGCCATCCACCTGCTC 1011
QY 847 CTGGGGGATGATGGGCACGTGAAGATCGCGAGCTTTGGCGTCAGCAACACAGTTTGAGGG 906
Db 1012 CTGGGGGATGATGGGCACGTGAAGATCGCGAGCTTTGGCGTCAGCAACACAGTTTGAGGG 1071
QY 907 AAGGAGCTCAGTGTGACGACGGGGGAAACCCAGCAATTCATGCGCCCGAGGCCATT 966
Db 1072 AAGGAGCTCAGTGTGACGACGGGGGAAACCCAGCAATTCATGCGCCCGAGGCCATT 1131
QY 967 TCTGATTCGGGCAGAGCTTCAGTGGGAAGCCTTCGATGTATGGGCCATCTGCGCTCAAG 1026
Db 1132 TCTGATTCGGGCAGAGCTTCAGTGGGAAGCCTTCGATGTATGGGCCATCTGCGCTCAAG 1191
QY 1027 TTGTAAGTCTTTGCTATGGGAAGTCCCATTCATCGAGATTCATCTGCGCTCCAC 1086
Db 1192 TTGTAAGTCTTTGCTATGGGAAGTCCCATTCATCGAGATTCATCTGCGCTCCAC 1251
QY 1087 AGGAGATCAAGAATGAGCCCGTGTGTTCTGAGGAGCCAGAAATCAGCGAGGAGTC 1146
Db 1252 AGGAGATCAAGAATGAGCCCGTGTGTTCTGAGGAGCCAGAAATCAGCGAGGAGTC 1311
QY 1147 AAGGACCTGATCTGAAGATGTTAGACAAGAAATCCCGAGACGAGAAATTCGGGTGCCAGAC 1206
Db 1312 AAGGACCTGATCTGAAGATGTTAGACAAGAAATCCCGAGACGAGAAATTCGGGTGCCAGAC 1371
QY 1207 ATCAAGTTGACCTTTGGGTGACCAAGAACGGGAGGAGCCCTTCCTTCGAGGAGGAG 1266
Db 1372 ATCAAGTTGACCTTTGGGTGACCAAGAACGGGAGGAGCCCTTCCTTCGAGGAGGAG 1431
QY 1267 CACTGACGCTGCTGAGGTGACAGAGGGGAGGTTAAGAACTCAGTCAGGCTCATCCCC 1326
Db 1432 CACTGACGCTGCTGAGGTGACAGAGGGGAGGTTAAGAACTCAGTCAGGCTCATCCCC 1491
QY 1327 AGCTGACCAACGCTGATCCTGCTGAAGTCCATGCTGAGGAAGGTTTCCTTTGGGAACCG 1386
Db 1492 AGCTGACCAACGCTGATCCTGCTGAAGTCCATGCTGAGGAAGGTTTCCTTTGGGAACCG 1551
QY 1387 TTTGAGCCCCAGGAGGAGGAGGATCGATCGATCGATCGATCGATCGATCGATCGATCG 1446
Db 1552 TTTGAGCCCCAGGAGGAGGAGGATCGATCGATCGATCGATCGATCGATCGATCGATCG 1611
QY 1447 AAGAAGGGTTTGTGAAGGGGCAAGAGCCAGAGCTCCCGCGCTCCAGGAAGAGGAG 1506
Db 1612 AAGAAGGGTTTGTGAAGGGGCAAGAGCCAGAGCTCCCGCGCTCCAGGAAGAGGAG 1671
QY 1507 GCTGCATCC 1515
Db 1672 GCTGCATCC 1680
```

RESULT 13

```
US-10-316-124-7
; Sequence 7, Application US/10316124
; Publication No. US20030152574A1
; GENERAL INFORMATION:
; APPLICANT: Logan, Thomas Joseph
; APPLICANT: Chun, Miyoung
; TITLE OF INVENTION: METHODS AND COMPOSITIONS TO TREAT
; FILE REFERENCE: CARDIOVASCULAR DISEASE USING 1419, 58765, AND 2210
; CURRENT APPLICATION NUMBER: US/10/316,124
; CURRENT FILING DATE: 2002-12-10
; PRIOR APPLICATION NUMBER: 60/339,995
; PRIOR FILING DATE: 2001-12-10
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 4427
; TYPE: DNA
; ORGANISM: Homo sapiens
```

```
US-10-316-124-7
Query Match 43.0%; Score 651.2; DB 15; Length 4427;
Best Local Similarity 71.3%; Pred. No. 4.3e-181;
Matches 873; Conservative 0; Mismatches 348; Indels 3; Gaps 1;
QY 267 GCTTATATGACAGGGGCTGCCAGCCACATCTCCCCCGGGCTGCGGGAGGCCACCAT 326
Db 360 GCTTATATGACAGGGGCTGCCAGCCACATCTCCCCCGGGCTGCGGGAGGCCACCAT 419
QY 327 CGAGTCCCCACCGTGGCCATCTCAGATCAGAGGACTCGTGCAGCTGAACACAGTACAA 386
Db 420 GGAGTCTCACCGTCTCCATCAGCGGTATGACGAGTCTGTGTCAGTGAATCAGTATAC 479
QY 387 GCTGAGATGAGATGGCAAGGTGCTTACGGTGTGTTGAGGTGGCCCTACAAAGAAAG 446
Db 480 CTTGAGGATGAATTTGGAAAGGGCTCTTATGTGTGTCGTAAGTTGGCTTACAAATGAAA 539
QY 447 TGAAGACACACACTATGCAATGAAAGTCTTTCCAAAAAGAAAGTTACTGAAGCAGTATGG 506
Db 540 TGAAGATACCTACTATGCAATGAAAGTCTTTCCAAAAAGAAAGTCTATCGGAGGCCGG 599
QY 507 CTTTCCACGTCGCCCTCCCCCGAGAGGGTCCAGAGCTGCCAGGAGGAGGAGGAGGAGGAG 566
Db 600 CTTTCCACGTCGCCCTCCCCCGAGGACCCCGGCGAGCTCTGGAGGCTGCATCCAGCC 659
QY 567 GCTGCTGCCCCGAGGAGCGGTGTACAGGAGATTTGCCATCTCTGAAGAGTGCACCGT 626
Db 660 CAGGGGCCCCATTTGAGCAGGTGTACAGGAAATTTGCCATCTCTGAAGAGTGCACCGC 719
QY 627 GAATGTGGTCAAACTGATCGAGGTCTGTGATCACCAGCTGAGGACCACTCTATTGTTGT 686
Db 720 CAATGTGTGAAGCTGTGTGAGGTCTGTGATGACCCCAATGAGGACCATCTGTACATGT 779
QY 687 GTTTCACCTCTCTGAGAAAGGGGCCGTCTATGGAAGTGCCTGTGACAAAGCCCTTCTCGGA 746
Db 780 GTTCGAACCTGGTCAACCAAGGGCCGTGTATGGAAGTGCACCCCTCAAAACCACTCTCTGA 839
QY 747 GAGCAAGCTCGCTCTACCTCGGGGACGTCTATCTTGGGCTCGAGTACTTGCATGCGCA 806
Db 840 AGACAGGCCCCCTTCTACTTCCAGGATCTGATCAAGGACATCGAGTACTTACACTACCA 899
QY 807 GAAGATCTGTCCACAGGACATCAAGCCATCCAACTGCTCTCTGGGGGATGATGGGCACGT 866
Db 900 GAAGATCTCACACCGTGTGATCAAACTCTCAACCTCTCTGTCGGAGAGATGGGCACAT 959
QY 867 GAAGATCGCGGCTTTGGCGTCAGCAACGAGTTGAGGGGAAACGAGCTGACGTGTCCAG 926
Db 960 CAAAGATCGTCACTTTGGTGTGAGCAATGAATTCAGGGGAGTGCAGCGCTCTCTCCAA 1019
QY 927 CAGCGGGGAACCCAGGATTCATGCCCCCGAGGCCATTTCTGATTCGGGCGAGGCTT 986
Db 1020 CACCGTGGGACGCCCCGCTTCATGGCACCCGAGTGCCTCTGAGACCCGCAAGATCTT 1079
QY 987 CAGTGGGAAGGCTTTGGATGTATGGGCCACATGCGGCTCAAGTTGTACTGTTGTCTATGG 1046
Db 1080 CTCGGAAGGCTTTGGATGTTTGGCCATCGGTGTGACACTATCTACTCTTGTCTTTGG 1139
QY 1047 GAAATGCCCATTCATCGAGATTTTCATCTGTGCCCTCCACAGAGAGATCAAGAAATGAGCC 1106
Db 1140 CCAATGTCCTTTCATGGACGAGGATCATGTGTTTACACAGTAAGATCAAGAGTCAGGC 1199
QY 1107 CGTGTGTTTCTCTGAGGAGCCAGAAATCAGCGAGGAGCTCAAGGACCTGATCCTGAAGAT 1166
Db 1200 CTTGGAATTTCCAGACCGCCGACATAGCTGAGGACTTTGAAGGACCTGATCACCCTGAT 1259
QY 1167 GTTAGCAAGAAATCCGAGAGAGAAATTTGGGTGCCAGACATCAAGTTGACCCCTTGGGT 1226
Db 1260 GCTGACAAAGAACCCCGAGTGTGAGGATCGTGTGTGCGGAAATCAAGTGCACCCCTGGGT 1319
QY 1227 GACCAAGAACGGGAGGAGGCCCTTCTCTCGAGAGGAGGAGCTGACGAGGTGGTGGAGT 1286
Db 1320 CACGAGGATGGGCGGAGCGCTTCCCGTTCGAGGATGAGAACTGCAAGCTGGTGGAAAT 1379
```



```
; APPLICANT: Zhao, Xumei
; APPLICANT: Anderson, Dustin
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; FILE REFERENCE: MRI-044
; CURRENT APPLICATION NUMBER: US/10/205,823
; PRIOR FILING DATE: 2002-07-25
; PRIOR APPLICATION NUMBER: 60/307,982
; PRIOR FILING DATE: 2001-07-25
; PRIOR APPLICATION NUMBER: 60/314,356
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/325,020
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: 60/341,746
; PRIOR FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: 60/362,158
; PRIOR FILING DATE: 2002-03-05
; NUMBER OF SEQ ID NOS: 455
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 51
; LENGTH: 4942
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-205-823-51

Query Match      43.0%; Score 651.2; DB 15; Length 4942;
Best Local Similarity 71.3%; Pred. No. 4.5e-181;
Matches 873; Conservative 0; Mismatches 348; Indels 3; Gaps 1;

QY 267 GCGTTATGCGACGGGCGCTGCGAGCCACATCTCCCCCGGGCGCTGGCGGAGGCCACCAT 326
DB 536 GCCCTACTCACCGTCAGCTCCCGCAGTCCTCGCTCGCTGCGCCCGCGCGCGACAGT 595
QY 327 CGAGTCCACACCGTGGCGCATCTCAGATCGAGAGGACTCGGTGCGAGTGAACAGTACAA 386
DB 596 GGAGTCTCACACCGTCTCCATCACGGGTATGCGAGGACTGTGTCAGCTGAATCAGTATAC 655
QY 387 GCTGCAGAGTGAATGGCAAGGTCCTACCGTGTGGTGGAGTGGCTGAGCTGCGCTACAAAG 446
DB 656 CTGGAAGATGAATTTGGAAGGGCTCTTATGGTGTGCTCAAGTTGGCTTACAATGAAA 715
QY 447 TGAAGACAGACACTATGCAATGAAAGTCCCTTTCCAAAAAGAAAGTTACTGAAAGCAGTATGG 506
DB 716 TGACAATACCTACTATGCAATGAAGTGTGTGTCCTCAAAAGAAAGTCAATCGGAGCGCG 775
QY 507 CTTTCCACGTGCGCTCCCGGAGAGGTCGCCAGGTCGCCAGGAGGAGGAGGAGCCAGCA 566
DB 776 CTTTCCACGTGCGCTCCCGGAGGTCGCCAGGTCGCCAGGAGGAGGAGGAGGAGGAGGAG 835
QY 567 GCTGCTGCCCTCGGAGCGGTGTACCGAGAGATTGCCATCTGAAAGAGTGGACCAAGT 626
DB 836 CAGGGGCCCCATTGAGCAGGTGTACCGAAGATTGCCATCTCAAGAGCTGGACCAAGT 895
QY 627 GAATGTGTCAAACTGATCGAGTCTCGATGATGATGATGATGATGATGATGATGATGATGAT 686
DB 896 CAATGTGTGAAGCTGCTGAGGTCTCGATGATGATGATGATGATGATGATGATGATGATGAT 955
QY 687 GTTGTGACCTCTCAGAAAGGGGCGCTCATGGAAGTGGCTGTGCAAGCCCTTCTCGGA 746
DB 956 GTTGAAGTGGTCAACCAAGGGGCGGTGATGGAAGTGGCCACCCCTCAAAACCACTCTCTGA 1015
QY 747 GGAGCAAGCTCGCTCTACTCGGGGAGCGTCATCTCTGGGCGCTCGAGTCTGCACTGCGCA 806
DB 1016 AGACCAAGGCGGCTTCTACTTCCAGGATCTGATCAAGAGCATCGAGTACTTACACTACCA 1075
QY 807 GAAGATCGTCCACAGGAGCATCAAGCCATCCAACTGCTGCTGGGGGATGATGGGCAAGT 866
DB 1076 GAAGATCATCCACCGTGAATCAAACTTCCAACTCTCTGGTGGGAGAGATGGGCAAT 1135
QY 867 GAAGATCGCGGACTTGGCGGTGAGCAACCAAGTTTGAAGGGGAGGAGGAGGAGGAGGAG 926
DB 1136 CAAGATCGCTGACTTGGTGTGAGCAATGAATTCAGGGGAGTGAACGCGCTCTCTCCAA 1195
```

Search completed: July 25, 2004, 14:06:56
Job time : 994 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 12, 2004, 04:23:48 ; Search time 57 Seconds
(without alignments)
852.223 Million cell updates/sec

Title: US-10-690-617-2

Perfect score: 2634

Sequence: 1 MEGGPAVCCQDPRAELVERV.....FEGGKSPELPGVQEDRAAS 505

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 78:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2452	93.1	505	2 A57156	Ca2+/calmodulin-de
2	1494	56.7	587	2 JCS669	Ca2+/calmodulin-de
3	1472.5	55.9	417	2 JEO191	calcium/calmodulin
4	1043	39.6	357	2 T37317	probable Ca2+/calm
5	620	23.5	652	2 S38666	serine/threonine-p
6	566	21.5	1142	2 S50632	protein kinase PKA
7	512	19.4	382	2 T47464	serine/threonine-p
8	507.5	19.3	560	2 S57252	probable serine/th
9	468	17.8	1246	2 T51085	related to protein
10	448	17.0	440	2 T14736	probable serine/th
11	437	16.6	440	2 T14735	probable serine/th
12	436.5	16.6	511	1 A56009	serine/threonine-s
13	434	16.5	461	2 T14822	probable serine/th
14	429.5	16.3	512	2 T07788	probable serine/th
15	426.5	16.2	512	1 JC1446	serine/threonine-s
16	423.5	16.1	504	2 T10449	probable serine/th
17	418.5	15.9	472	2 B90100	SNP-related kinase
18	416.5	15.8	441	2 E85362	hypothetical prote
19	415.5	15.8	441	2 C84667	probable protein k
20	413.5	15.7	435	2 E84707	probable protein k
21	412.5	15.7	512	2 T52633	serine/threonine-s
22	410	15.6	389	2 S52242	protein kinase (EC
23	409.5	15.5	1192	2 T18611	probable serine/th
24	409	15.5	445	2 T50802	serine/threonine p
25	406.5	15.4	469	2 B84644	probable protein k
26	406	15.4	480	2 A86427	probable serine/th
27	405.5	15.4	504	2 T07415	probable serine/th
28	403.5	15.3	633	1 A26030	serine/threonine-s
29	403	15.3	576	2 T41587	probable carbon ca

GIN4 protein - yea
serine/threonine-s
protein H39E23.1
probable serine/th
probable serine/th
FOG2 protein - yea
protein T25K16.13
myosin-light-chain
probable protein k
Ca2+/calmodulin-de
hypothetical prote
protein kinase (EC
hypothetical prote
p69Bg3 protein - A
SNF1-related prote
serine/threonine-s

ALIGNMENTS

RESULT 1

A57156

Ca2+/calmodulin-dependent protein kinase (EC 2.7.1.123) IV - rat

N:Alternate names: Cam-kinase kinase; microtubule-associated protein 2 kinase

C:Species: Rattus norvegicus (Norway rat)

C>Date: 23-Feb-1996 #sequence_revision 23-Feb-1996 #text_change 21-Jul-2000

C:Accession: A57156; J04789

R:Tokumitsu, H.; Emslen, H.; Soderling, T.R.

J. Biol. Chem. 270, 19320-19324, 1995

A:Title: Characterization of a Ca(2+)/calmodulin-dependent protein kinase cascade. Molec

A:Reference number: A57156; PMID:95370263; PMID:7642608

A:Accession: A57156

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-505 <TK>

A:Cross-references: GB:I42810; NID:g986940; PIDN:AAC42070.1; PID:g986941

R:Okuno, S.; Kitani, T.; Fujisawa, H.

J. Biochem. 119, 1176-1181, 1996

A:Title: Evidence for the existence of Ca2+/calmodulin-dependent protein kinase IV kinase

A:Reference number: JC4789; PMID:96425004; PMID:8827455

A:Accession: JC4789

A:Molecule type: mRNA

A:Residues: 1-44, 'P', '46-80, 'G', '82-505 <OKU>

A:Cross-references: GB:AB023658; NID:g4512333; PIDN:BAA75246.1; PID:g4512334

A:Experimental source: brain

C:Comment: This enzyme is a Ca2+-responsive multifunctional protein kinase. It plays rol

C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolo

C:Keywords: ATP; brain; calmodulin binding; phosphotransferase; protein kinase

F:126-409/Domain: protein kinase homology <KIN>

F:134-142/Region: protein kinase ATP-binding motif

Query Match 93.1%; Score 2452; DB 2; Length 505;

Best Local Similarity 93.1%; Pred. No. 3 3e-101;

Matches 470; Conservative 11; Mismatches 24; Indels 0; Gaps 0;

QY 1 MEGGPAVCCQDPRAELVERVAIDVTHLEADGGPEPTRNGVDPPPPARAASVIGPSTGR 60

DB 1 MERSPAVCCQDPRAELVERVAISVAHLEAEAGPEPASNGVDPPSARAASVIGPSASR 60

QY 61 LIPAPPSLSARKLSQERPAQGYATGASHISPRAWRPRTIESHHVAISDAE 120

DB 61 PTFVPSLSARKFSQERPAESCLEAQGPYSTGASHMSPPAWRPRTIESHHVAISDTE 120

QY 121 DCVQLNQYKIQSEIGKGVVRLAYNSEDHRYAMKVLKKKLLKQYGFPRPPRPGSQ 180

DB 121 DCVQLNQYKIQSEIGKGVVRLAYNREDHRYAMKVLKKKLLKQYGFPRPPRPGSQ 180

QY 181 AAQGGPAKOLLFLERVYQIEAILKKLDHVNKKLIEVLDDPAEDNLVLFVLLRKGPMVE 240

DB 181 APQGGPAKOLLFLERVYQIEAILKKLDHVNKKLIEVLDDPAEDNLVLFVLLRKGPMVE 240

QY 241 VPCDFPFSEQARLYLRDVLGVLEHQCQKIVHRDIKPSNLLGGDGHVKIADFGVSNQF 300

Db 241 VPCDKPPEQARLYLRDIILGLVLCQKIVHRDIKPSNLLDGDGHVKIADFGVSNQF 300
Qy 301 EGNDAQLSSTAGTAPFAPEAISDSGQSFSGKALDVWATGVTLYCFVYGKCFDIDFTILA 360
Db 301 EGNDAQLSSTAGTAPFAPEAISDTGQSFSGKALDVWATGVTLYCFVYGKCFDIDFTILA 360
Qy 361 LHRKIKNEPVFPPEPEISEELKDLILKMLDKNPETRIGVDPDIKILHPWTKNGEPELPSE 420
Db 361 LHRKIKNEAVFPPEPEISEELKDLILKMLDKNPETRIGVSDIKILHPWTKHGEPELPSE 420
Qy 421 BEHCSVVVEGEVKNVRLIPSWTIVLVKSMRKRSFGNPFEPQARREERSMSAPGNL 480
Db 421 BEHCSVVVEEENSVKLLIPSWTIVLVKSMRKRSFGNPFEPQARREERSMSAPGNL 480
Qy 481 LVKEGFGEGGKSPGLPGVQDEAAS 505
Db 481 LLKEGCGEGGKSPGLPGVQDEAAS 505

RESULT 2

JC5669
Ca2+/calmodulin-dependent protein kinase kinase (EC 2.7.1.1-) beta chain - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 20-Nov-1997 #sequence_revision 20-Nov-1997 #text_change 24-Sep-1999
C:Accession: JC5669; PC4493
R:Kitani, T.; Okuno, S.; Fujisawa, H.
J. Biochem. 122, 243-250, 1997
A:Title: Molecular cloning of Ca2+/calmodulin-dependent protein kinase beta.
A:Reference number: JC5669; MUID:97420710; PMID:9276695
A:Accession: JC5669
A:Molecule type: mRNA
A:Residues: 1-587 <KIT>
A:Cross-references: GB:AB018081; NID:g3702720; PIDN:BAA33524.1; PID:d1034490; PID:g3702720
A:Experimental source: cerebellar
A:Accession: PC4493
A:Molecule type: protein
A:Residues: 425-501 <KI2>
A:Experimental source: cerebellar
C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolog
C:Keywords: ATP; phosphotransferase
F:162-445/Domain: protein kinase homology <KIN>
F:170-194/Domain: ATP-binding #status predicted <ATP>
F:480-493/Domain: calmodulin-binding #status predicted <CAB>

Query Match 56.7%; Score 1494; DB 2; Length 587;
Best Local Similarity 64.5%; Pred. No. 4.7e-59;
Matches 283; Conservative 75; Mismatches 55; Indels 26; Gaps 3;

Qy 64 ARPSLSARKLSQRPAGSVLEAAGP-----YATGPASHISPRAW 105
Db 89 SRSLSGKMSLQR-----SOGGPASSSLDMNGRCICPSLSYSPASSPQSPPMR 141
Qy 106 RPTIESHHVAISDAEDCVQLNQYKLSQEIIGKYGAYGVVRLAYNESDRHYAMKVLKSKLL 165
Db 142 RPTVESHVHSITGLQDCVQLNQYTLKDEIGKSGYGVVKLAYNENDNTYAMKVLKSKLLI 201
Qy 166 KQYGFPPRRPRGSAQAGGPAKQLPLERVYQIAILKLDHVNVLKLEVLDDPAEDN 225
Db 202 RQAGFPPRRPRGRTRPAPGGCIQPRGTEQYQIAILKLDHFNVLKLEVLDDPDNEH 261
Qy 226 LYLVDLLRKGPVMEVPCDKPFSBEQARLYLRDVLGLVLCQKIVHRDIKPSNLLDGD 285
Db 262 LYMVVELVNGQPVMEVTLKPLSEDAQRFYFDLILKGLVLYHQIKIHRDIKPSNLLVGE 321
Qy 286 DGHVKIADFGVSNQFEGNDAQLSSTAGTAPFAPEAISDSGQSFSGKALDVWATGVTLYC 345
Db 322 DGHKIADFGVSNFEGKSDALLSNTVGTAPFAPEAISLSETRKIFSGKALDVWAMGVTLYC 381
Qy 346 FVYGKCFDIDFTILALHRKIKNEPVFPPEPEISEELKDLILKMLDKNPETRIGVDPDIK 405
Db 382 FVFGCCPFMDRIMCLHSKISQALEFPDQDIAEDLKDLITRMLDKNPESRIVVPEIKL 441

Qy 406 HPWTKNGEPLPSEBEHCSVVVEGEVKNVRLIPSWTIVLVKSMRKRSFGNPFEP 465
Db 442 HPWTRREGAELPSEDENCTLVETBEVENSVKHIPSLATVILVKTMRKRSFGNPFEP- 500
Qy 466 QARREERSMSAPGNLLVKE 484
Db 501 GSRREERSLSAPGNLLTKK 519

RESULT 3

JE0191
calcium/calmodulin-dependent protein kinase kinase (EC 2.7.1.1-) - human
X:Alternate names: CaMKK
C:Species: Homo sapiens (man)
C:Date: 21-Aug-1998 #sequence_revision 21-Aug-1998 #text_change 16-Dec-1998
C:Accession: JE0191
R:Hsu, L.S.; Tsou, A.P.; Chi, C.W.; Lee, C.H.; Chen, J.Y.
Biomed. Sci. 5, 141-149, 1998
A:Title: Cloning, expression and chromosomal localization of human Ca2+/calmodulin-depe
A:Reference number: JE0191
A:Accession: JE0191
A:Molecule type: mRNA
A:Residues: 1-417 <HSU>
C:Comment: This enzyme functions as a key intermediaries in Ca2+/calmodulin-driven sign
C:Genetics:
A:Gene: CaMKK
A:Map position: 12
C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolog
C:Keywords: phosphotransferase
F:47-330/Domain: protein kinase homology <KIN>

Query Match 55.9%; Score 1472.5; DB 2; Length 417;
Best Local Similarity 68.9%; Pred. No. 3e-58;
Matches 272; Conservative 72; Mismatches 50; Indels 1; Gaps 1;

Qy 90 PYATGPASHISPRARRPTTESHVVAISDAEDCVQLNQYKLSQEIIGKYGAYGVVRLAYNES 149
Db 11 PYSVPSSPQSPLPRPRTVESHVSVITGMQDCVQLNQYTLKDEIGKSGYGVVKLAYNEN 70
Qy 150 EDRHYAMKVLKSKLLKQYGFPPRRPRGSAQAGGPAKQLPLERVYQIAILKLDH 209
Db 71 DNTYAMKVLKSKLLIRQAGFPFRPPRPRGTRPAPGGCIQPRGTEQYQIAILKLDHP 130
Qy 210 NVVKLIEVLDDPAEDNLVLDLARKGPVMEVPCDKPFSBEQARLYLRDVLGLVLCQ 269
Db 131 NVVKLIEVLDDPDNEHDLYMVVELVNGQPVMEVTLKPLSEDAQRFYFDLILKGLVLYHQ 190
Qy 270 KIVHRDIKPSNLLDGDGHVKIADFGVSNQFEGNDAQLSSTAGTAPFAPEAISDSGQSF 329
Db 191 KIIHRDIKPSNLLVGEHGHKIADFGVSNFEGKSDALLSNTVGTAPFAPEAISLSETRKIF 250
Qy 330 SGKALDVWATGVTLYCVYQKCPDIDFTILALHRKIKNEPVFPPEPEISEELKDLILKM 389
Db 251 SGKALDVWAMGVTLVCFVFGQCFPMDERIMCLHSKISQALEFPDQDIAEDLKDLITRM 310
Qy 390 LDKNPETRIGVDPDIKILHPWTKNGEPELPSEBEHCSVVVEGEVKNVRLIPSWTIVIL 449
Db 311 LDKNPESRIVVPEIKLHPWTRHGAELPSEDENCTLVETBEVENSVKHIPSLSATVIL 370
Qy 450 VKSMRKRSFGNPFEPQARREERSMSAPGNLLVKE 484
Db 371 VKTMRKRSFGNPFEP- GSRREERSLSAPGNLLTKK 404

RESULT 4

T37317
probable Ca2+/calmodulin-dependent protein kinase kinase (EC 2.7.1.1-) - Caenorhabditis
C:Species: Caenorhabditis elegans
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 21-Jul-2000
C:Accession: T37317
R:Edelman, A.M.; Mitchellhill, K.I.; Selbert, M.A.; Anderson, K.A.; Hook, S.S.; Stapleton
J. Biol. Chem. 271, 10806-10810, 1996
A:Title: Multiple Ca-calmodulin-dependent protein kinase kinases from rat brain.

Best Local Similarity 34.4%, Pred. No. 1.3e-20;	
Matches 168; Conservative 67; Mismatches 176; Indels 78; Gaps 14;	
Qy	52 SVIPGTSRLLPAPSPISARKLSQIER--PAGSYLEAQAGPYATG-----PASH- 98
Db	30 SEVPPSVFPEYVHKAIQKTSDFRKENYSAGDYVIAPLGEREGSSLTSHWTFQPKHN 89
Qy	99 -----ISPANRR-----PTTESHVIAISDAEDCVOLNQYKLOSBIGKAY 139
Db	90 QRLYSDFQAQRQWKQLQEWGEYKETKKIRKFRDFSGRK--YINHYBIILKELGRMH 14
Qy	140 GVVKLAYNESEDRHYAMKVLSSKKLLKQYGFRRPPRPGSQAQGGPAKQLLPLERYVQE 199
Db	147 GKVKLGRDVTRELLAIKIIPKE-----RRP-----XLRANASSQKEKVTRE 190
Qy	200 IAILKKLDHVNVLKIEVLDDPDAEDNLIYLVFDLLRKGFMVEVCPDKP-FSEEQARLYLRD 258
Db	191 IAILKKCVHNVKRLREVIDDPSSTKVYLVLYMSGGVEPWTDCDSVLSISEARQYFRD 250
Qy	259 VILGLEYLHCQKVHRDIPKSNLLIGDDGHVKIADFGVS-----NQEGNDAQLSSTAGT 313
Db	251 VWLGLEYLHYQGIHHRDIKANLLNSNCVKLSDFGVSYIANGLNEDNDVELAKTVGT 310
Qy	314 PAFMAPEAI---SDSGQSFSGKALDWATGVTLIYCFYVYKGCPIDDPILALHRKIKNEPV 370
Db	311 PAFPAPELCTDLDPRPKISEALDWALGVTLFCLLFGRCPFNASMEYELDKIVNERL 370
Qy	371 VFPEPPISELKDLILKMLDKNPETHRIGVDPDKLHPWTWNGEEP----LPSEEEHCSV 426
Db	371 NIPSTPDIGEGRDLLKRLCKOPEQRITIVVKLHPWTLTDGLKDPKFWLQNTDPTSVSR 430
Qy	427 VEVTGEGVKNSVRLIPSWTTVILVKMLRKRSFGNPF---EPQARRERSMSAPGNLLVVK 483
Db	431 VEVSTDEVASAI SLV-----GRLR-RKLGLKFRFRPKARVDFDSSSVSDSSIC 479
Qy	484 EGFGEKGKS 492
Db	480 RPESSGNSNS 488

RESULT 6

S50632 protein kinase PAK1 (EC 2.7.1.1-) - yeast (Saccharomycetes cerevisiae)

N:Alternate names: protein YER129w

C:Species: Saccharomycetes cerevisiae

C:Date: 28-Jan-1995 #sequence revision 12-May-1995 #text_change 23-Mar-2001

C:Accession: S50632; S47952; S43223

R:Dietrich, F.S.

submitted to the EMBL Data Library, December 1994

A:Description: The sequence of S. cerevisiae cosmid 9781, 8198, 9115, 9981, an

A:Reference number: S50437

A:Accession: S50632

A:Molecule type: DNA

A:Residues: 1-1142 <DIE>

A:Cross-references: EMBL:U18916; NID:g1384128; PIDN:AAC03227.1; PID:g603368; MI

R:Hovland, P.G.; Sclarani, R.A.

submitted to the EMBL Data Library, August 1994

A:Description: PAK1: DNA Polymerase Alpha Suppressing Protein Kinase.

A:Reference number: S47952

A:Accession: S47952

A:Molecule type: DNA

A:Residues: 1-170, 'H', 172-265, 'DS', 269-657, 'LD', <HOV>

A:Cross-references: EMBL:U03398; NID:g532797; PIDN:AAC49840.1; PID:g532798

R:Mulligan, J.T.; Dietrich, F.S.; Hennessy, K.M.; Sehl, P.; Komp, C.; Wei, Y.;

submitted to the EMBL Data Library, February 1993

A:Reference number: S30812

A:Accession: S43223

A:Molecule type: DNA

A:Residues: 1-176, 'VYQIK' <MUL>

A:Cross-references: EMBL:L11119

C:Genetics:

A:Gene: SGD:PAK1

A:Cross-references: SGD:S0000931; MIPS:YER129w

A;Map position: 5R
C;Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolog
C;Keywords: ATP; phosphotransferase; protein kinase
F;131-448/Domain: protein kinase homology <KIN>
F;139-147/Region: protein kinase ATP-binding motif

Query Match 21.5%; Score 566; DB 2; Length 1142;
Best Local Similarity 31.6%; Pred. No. 5.2e-18;
Matches 151; Conservative 82; Mismatches 135; Indels 110; Gaps 14;

QY 98 HISPRWRPTIESHHVAISDAEDCVQ-----LNQYKLOSEIGKAGY 140
DB 86 HISSLAKTPTTSSFCSSGSKNKVKETNRISLTYPVSKRVLTVEIILKELGHQGH 145
QY 141 VRLAYNESEDRHYAMKVLK--KKLLKQYGFPPRPPRGSQAAGGPAKQLLPLRVYQ 198
DB 146 KVLARDILSKQLVAIKIVDRHEKKQKFFFIK-----SSKISENDKIKR 191
QY 199 EIALKELKLDHNVVVKLEIVLDDPAEDNLYLVFLLRKGPVMEVPCD-----KPFSE 249
DB 192 EIALMKCHKHVVQLIEVLDDLSKRYLVLEYSRGEVVKWCPDCEMSDAKGPSLLSF 251
QY 250 EQARLYRDVILGLYHLCQKIVHRDIKPSNLLGDDGHVXIADFGVS-----NQPEGN 303
DB 252 QETREILRGVVLGLYHLYQYIIHRDIKPAMLISGDTVKISDFGVSLAASSTNSDS 311
QY 304 DA---OLSSTAGTPAFMAPE-AISDSG-----QSTSGKA---LDVWATGVTLYC 345
DB 312 ESDLDELAKTVGTPAFAPEMCGEDAFTRYNUTKENLFRGSCISFMIDVWAGVTLYC 371
QY 346 FVYGKCPFFIDFIALHRKIKNEVPFPEEPIIS-----EELKDLILKMLDK 392
DB 372 LLFCMLPFSSDFELKLFKEIVNDFLKPFTFKFQISQNVKSVCSCEEYEMAKDILLKLEK 431
QY 393 NPETRIGVPDIKLHPWT-----KNGEPLPS-----EEHSCSVETGSEV 434
DB 432 NPQKRMTPAIKHPFVSDFWDPHVDENKLLSVLEQLKRFQCNQDQFEPISISKHEL 491
QY 435 KNSVRLIPSTVTILVKSM-----LRKRSFGNPFEPQARREERSMSAPCNLLAVKG 485
DB 492 KNAVSGVKKIKESVLKSIPLKDPDSLKNKNYLHPTETTRGRGD-----ANVIVSEG 543

RESULT 7
T47464
serine/threonine-protein kinase-like protein - Arabidopsis thaliana
N;Alternate names: protein T14D3.180
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 20-Apr-2000
C;Accession: T47464
R;Jordan, N.; Baugert, S.; Wiedemann, R.; Voss, H.; Unseld, M.; Mewes, H.W.; Lemcke, K.
submitted to the Protein Sequence Database, February 2000
A;Reference number: 224467
A;Accession: T47464
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-382 <JOR>
A;Cross-references: EMBL:AL138649
A;Experimental source: cultivar Columbia; BAC clone T14D3
C;Genetics:
A;Map position: 3
A;Introns: 86/3; 111/3; 126/3; 151/3; 181/3; 222/3; 251/3; 312/3; 339/1
A;Note: T14D3.180

Query Match 19.4%; Score 512; DB 2; Length 382;
Best Local Similarity 37.0%; Pred. No. 4.4e-16;
Matches 119; Conservative 56; Mismatches 107; Indels 40; Gaps 9;

QY 107 PTIESHHVAISDAEDCVQ--LNQYKLOSEIGKAGYVRLAYNESEDRHYAMKVLKSKLL 165
DB 75 PVKETNKLTRGEDGDKNTNEFVRERKIGSGYGVLYRSTVDDKHVYAKAFHKSLS 134
QY 166 KQYGFPPRPPRGSQAAGGPAKQLLPLRVYQEIALLKLDHNVVVKLEIVLDDPAEDN 225

Db 135 RLRVAP-----SETMGD-----VLRWIMKTLHEPNVNLIEVIDDPEFDD 177
QY 226 LYLVPDLL-----RKGPVMEVPCDKPFSEQARLYLRDVILGLYHLCQKIVHRDIKP 278
DB 178 FYMVLVYDVGKWAYDDSGP-----PGALGEITARKYLRDVVVAGLVYLHAHNVHGDIKP 231
QY 279 SNLLLGDDGHVXIADFGVSQFEGNDAQLSSTAGTAPFMAPEALISDSQSFSKALDVA 338
DB 232 DNLAVTSTGRVKGIDFVS---ODDDQLRRSPGTPVFTAPECC--LGITYSGRSADTWA 286
QY 339 TGVTLYCFVYKCKCFIDDFIALHRKIKNEVPFPEEPISEELKDLILKMLDKNPETRI 398
DB 287 VGVTLYCMILGOYFPLGDTQDYDKIVHNPILIPE--GLNPLRLDIEGLLCKDPQORM 344
QY 399 GVPDIKLHPWTVKNGEPLPSE 420
DB 345 TLKAVAEHPWIT--GEDGAISE 364

RESULT 8
S57252
probable serine/threonine-specific protein kinase (EC 2.7.1.1-) YGL179c - yeast (Sacchar
N;Alternate names: protein BIE560; protein G1618
C;Species: Saccharomyces cerevisiae
C;Date: 28-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 19-Apr-2002
C;Accession: S57252; S64196; S56168
R;Coglievina, M.; Bertani, I.; Klima, R.; Zaccaria, P.; Bruschi, C.V.
Yeast 11, 767-774, 1995
A;Title: The DNA sequence of a 7941 bp fragment of the left arm of chromosome VII of Sa
ation and a putative serine/threonine protein kinase gene.
A;Reference number: S57252; MUID:95397594; PMID:7668046
A;Accession: S57252
A;Status: nucleic acid sequence not shown
A;Molecule type: DNA
A;Residues: 1-560 <COG>
A;Cross-references: EMBL:X83690; NID:g794143; PIDN:CAA58659.1; PID:g794144
R;Bruschi, C.V.; Coglievina, M.; Bertani, I.; Klima, R.; Zaccaria, P.; Delneri, D.
submitted to the Protein Sequence Database, May 1996
A;Reference number: S64183
A;Accession: S64196
A;Molecule type: DNA
A;Residues: 1-560 <BRU>
A;Cross-references: EMBL:Z72701; NID:g1322788; PIDN:CAA96891.1; PID:e243787; PID:g13227
A;Experimental source: strain S288C
C;Genetics:
A;Gene: SGD:TOS3
A;Cross-references: SGD:S0003147
A;Map position: 7L
A;Note: YGL179c
C;Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolog
C;Keywords: ATP; phosphotransferase; serine/threonine-specific protein kinase; transmem
F;48-344/Domain: protein kinase homology <KIN>
F;56-64/Region: protein kinase ATP-binding motif
F;256-272/Domain: transmembrane #status predicted <TM>

Query Match 19.3%; Score 507.5; DB 2; Length 560;
Best Local Similarity 35.2%; Pred. No. 1e-15;
Matches 128; Conservative 58; Mismatches 115; Indels 63; Gaps 10;

QY 125 LNQYKLOSEIGKAGYVRLAYNESEDRHYAMKVLKSKLLKQYGFPPRPPRGSQAAG 184
DB 47 LNNFEILATLNGQYGVKVLARDLGTGALVAIKILNR-----FEKR-----S 88
QY 185 GPAKQL-LPLERVYQEIALLKLDHNVVVKLEIVLDDPAEDNLYLVFLLRKGPV----- 238
DB 89 GYSLQLKVENPRVNOQIEVMKRCHEHNVVELYILNDPESTKVLVLEYSRGEVFKWCP 148
QY 239 --NEVPCKDP--FSEQARLYLRDVILGLYHLCQKIVHRDIKPSNLLGDDGHVXIAD 294
DB 149 NKMEIKAVGSEILTFQOSRKVLDVVGLEYLHSGIHTHRDIKPSNLLISNGTVKLIS 208
QY 295 GV-----SNOPEGNDAQL--SSTAGTAPFMAPEALISDSQSFSKALDVAWATGV 345

Db 209 GVAMSTATGNTIOSSHEQLKSRALGTPAPFAPELCSSTEXEYSCSSAIDWISLGVIIYC 268
Qy 346 FVYKCPDIDFIALHHRKIKNEPVREPEEISE-----ELKDLILKMLDKN 393
Db 269 LFLKLPNAGSLELFIINIKPLFEFFSYEMLNATGATIMEYTDKALLKLLQKD 328
Qy 394 PETRIGVPDIKLHPVTKNGBEPLPS-----BEEH-----CSVVEVTEGEVNSVR 439
Db 329 PDKRIKLADIKVHPFMCHYGKSDAASVLTNLTETHELKVSPPSSCKRVELSLPNSSFA 388
Qy 440 LIPS 443
Db 389 SLDS 392

RESULT 9
T51085
related to protein kinase PAK1 [imported] - Neurospora crassa
N;Alternate names: protein B2A19.190
C;Species: Neurospora crassa
C;Date: 21-Jul-2000 #sequence_revision 21-Jul-2000 #text_change 21-Jul-2000
C;Accession: T51085
R;Schulze, U.; Aign, V.; Hoheisel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura,
submitted to the Protein Sequence Database, July 2000
A;Reference number: Z25286
A;Accession: T51085
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1246 <SCH>
A;Cross-references: EMBL:AL390092; GSPDB:GN00116; NCSP:B2A19.190
A;Experimental source: BAC clone B2A19; strain OR74A
C;Genetics:
A;Gene: NCSP:B2A19.190
A;Map position: 6
A;Introns: 95/2

Query Match 17.8%; Score 468; DB 2; Length 1246;
Best Local Similarity 21.5%; Pred. No. 1.1e-13;
Matches 165; Conservative 91; Mismatches 168; Indels 344; Gaps 21;

Qy 37 PTRNGVDP-----PPRAAASVIPGSTSLPLA-----RPSLSARKLSLQERPAAGSVLE 85
Db 19 PTQHHPRELHPPLPP---MASQIPHTANSTPVSSPGLFSPISPRNMSLSQSPSE--- 72
Qy 86 AQAGPYATGPASHIPRAWRRPTIESHHVAIDAEQCVQ-----LNQYKLOSEIGKG 137
Db 73 -NTTAPTLHSPYLP-----LQSHKVRTHKANVEHDYTTGRKAINQYHIIIEELGEG 124
Qy 138 YGVVRLAYNESEDRHYAMKVLKLLKQYGFRRPPRGSAQGGPAKQLLPLRVY 197
Db 125 MHGKVKLARNTQTGDNAVKII-----PRFSKRLRGKVTAMSTQDKSK 168
Qy 198 QEIALKLLDHNVVKLLEVLDDPAEDNLVLYFD----- 231
Db 169 REIALKLTIRHPNVALLIEDDPELKYTMVLEHVELGVVVRKKGPHICAYERRQEQ 228
Qy 232 ----- 231
Db 229 REQLGALPDAREEYVLFMEKROAEKAKRAHVARQAHNNAEFWLSLEFGAADDLDDPH 288
Qy 232 ----- 231
Db 289 ARSLGRDSSIPFAGDWLGLSRVTSRAPRTSQSMKSISSRNTPOPSEPDPHASIASVVE 348
Qy 232 -----LLRKGPMVE----- 240
Db 349 VDEDEMETPHGSHINQDSAIEDSLFAPRENHSLFKRSPSMADSLSHMSSVDYNRVHD 408
Qy 241 -----VPCDKPFSEEARLYLRDVLIGLEYLHCQKIVHRDIKPSNLLDGDGHVKT 292
Db 409 PFVDDYSYVFC---FTIDKARSAFRDVLGLLEVLHVEGVVHRDIKPNALLCTKDRVKIS 465

Qy 293 DFGVS-----NOFEGNDAQISSTAGTAPWAPDE-AISDSQSQSPSG 331
Db 466 DFGVSYFGRPLRDGEDPEVSESEARDFD-NDLEAKTAVGTAPFAPELCYTDTYDDRP 524
Qy 332 -----KALDQWATGVTLVCFVYCKCPFDDFIALHHRKIKNEPVREPEE 373
Db 525 QQPKITQIDVMSLUGVTLYCLIFARIPLADEWRMFKKIATEDIYIPRQRLRPVDPSTK 584
Qy 374 -----EEP-----EISEELKDLILKMLDKNPKETRIGVPDIKLHPWTKN 412
Db 585 PDEKSLYTRVNRDPYRNDDDEPLVEEDNDLYDLLSKMLTKNPEKRIRLRDVKRHPWLKD 644
Qy 413 GEEPL-----PSEEHCSVVVTEGEVNSVRLIP-----SWTIVLVKSMLEKRSFGN 461
Db 645 IDNVIAWLDDTDPSSRTAGRIQVDRDITQAV--VPLTFVERAKLAFKKTIVTKFTHRGD 702
Qy 462 PPEQARREERSMSAPGNLLVKEGEGGKPEL-----PGVQDEEAAAS 505
Db 703 RSESVSSKRKRAITSAAS-----SAESPAGVGIPTFGVDRGERKS 741

RESULT 10
T14736
probable serine/threonine kinase (EC 2.7.1.1-) SNFL2 - sorghum
N;Alternate names: SNF1 protein kinase homolog SNFL2
C;Species: Sorghum bicolor (sorghum)
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 21-Jul-2000
C;Accession: T14736
R;Amnen, F.; Stockhaus, J.
Plant Mol. Biol. 36, 529-539, 1998
A;Title: Characterization of a Sorghum bicolor gene family encoding putative protein kin.
A;Reference number: Z18177; MUID:98145442; PMID:9484448
A;Accession: T14736
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-440 <ANN>
A;Cross-references: EMBL:Y12465; NID:g2632253; PIDN:CAA73068.1; PID:g2632254
A;Experimental source: cultivar TX 430
C;Genetics:
A;Gene: SNFL2
C;Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolo
C;Keywords: phosphotransferase; serine/threonine-specific protein kinase

Query Match 17.0%; Score 448; DB 2; Length 440;
Best Local Similarity 29.5%; Pred. No. 3.3e-13;
Matches 119; Conservative 75; Mismatches 139; Indels 70; Gaps 11;

Qy 124 QLNQYKLOSEIGKGAYGVVRLAYNESEDRHYAMKVLKLLKQYGFRRPPRGSAQ 183
Db 9 RVGKVELGRTIGEGTFAKVRFAKTETGEPVAKILDRKVKLRH----- 52
Qy 184 GCPAKQLLPLERVYQEIAILKKLDHNVVVKLIEVLDDPAEDNLVLYFDLLRKGPMVYPC 243
Db 53 -----KWEQIKREISTMKLTIKHPNVVRIVEVWG--SKTKIYIVLEVVTGGELFDITIA 103
Qy 244 DK-PPSEQARLYLRDVLIGLEYLHCQKIVHRDIKPSNLLDGDGHVKADEGVS--NQF 300
Db 104 NHGRNREDEARRYFOQLINAVDYCHSRGVYHRDLAPENLLDSYGNLKVSDFGLSALSQ 163
Qy 301 EGNDQAQLSTAGTAPWAPAEISDSQSFGSKALDQWATGVTLVCFVYCKCPFIDDTILA 360
Db 164 IKDDGLLHTTCGTPNYVAPEVLEDDQ--YDGAMADLWSCGVILFVLLGYLDFEDSNLMT 221
Qy 361 LHRKIKNEPVVPEPEISEELKDLILKMLDKNPKETRIGVPDIKLHPWTKNKEEP---- 416
Db 222 LYKKISNAEYFP--PWTSPFAKRLLTFLDPNPWTRITIPLELDEWFKKGKPEFDE 279
Qy 417 -----LPSEEHCSVVEVTEGEVNSVRLIPSWTIVLVKSMLEKRSFGNPF 463
Db 280 KYDTPLDDVDVAFVNDSEEHVTEKKEEPVVLNAPELI-----SRSAGLNLGNLF 329
Qy 464 --EPQARREERSMS--APGNLLVKE-----GFCGEGGKSP 495

Db 330 DSEQEKKRETRFTSKCPKPKETVRKIEBAAKPLGRGVQKNYK 372

RESULT 11

Tl4735

N;Alternate names: SNF1 protein kinase homolog SNFL1
C;Species: Sorghum bicolor (sorghum)
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 21-Jul-2000
C;Accession: Tl4735

R;Annen, F.; Stockhaus, J.
Plant Mol. Biol. 36, 529-539, 1998

A;Title: Characterization of a Sorghum bicolor gene family encoding putative protein kinases
A;Reference number: Z18177; MUID:9814542; PMID:9484448
A;Accession: Tl4735

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-440 <ANN>

A;Cross-references: EMBL:Y12464; NID:g2632251; PIDN:CAA73067.1; PID:g2632252

A;Experimental source: cultivar TX 430

C;Genetics:

A;Gene: SNFL1

C;Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolog
C;Keywords: phosphotransferase; serine/threonine-specific protein kinase

Query Match 16.6%; Score 437; DB 2; Length 440;

Best Local Similarity 29.7%; Pred. No. 1e-12;

Matches 114; Conservative 72; Mismatches 136; Indels 62; Gaps 10;

QY 124 QLNQYKLOSEIGKAGYGVRLAYNESEDRHYAMKVLKSKLLKQYGFPPRPGSQAAQ 183

Db 9 RVGKYLGRITGEGTAKVRFARDTVTGEAVAKILDKVKLKH----- 52

QY 184 GGPAKQLPLERVYQETAILKLDHVNVLKLEVLDDPAEDNLVLFDLRKGPMVEVP 242

Db 53 -----KMEQKREISTMKLKHNVNRIEYVMG--SKTIYIVLEFATGELFQRIV 103

QY 243 CDKPFSEEQARLYLRDVLGLEYLHCQKIVHRDIKPSNLLLGDDGHVKIADFGVS--NQF 300

Db 104 NHGRMREDEARRYFQQLINAVDYCHSGVYHRDLKPENLLDSYGNLKVSDFGLSALSOQ 163

QY 301 EGNDAQLSSTAGTAPFAPWAPAI SDSQSGSKALDVWATGVTYCYFYGKCPIDDFILA 360

Db 164 MKDDGLLHTTCGTFTYVAPEVLEDOG--YDGMADLWSCGVILFVLLAGYLPEDSNLMT 221

QY 361 LHRKINPEVVPFEEPEISEELKDLILKMLDKNPETRIGVDPDKLHPWTKNGEPP----- 416

Db 222 LYKKISNAEFTFP--FWTSPPAKRLTLRLDPMTRITIPETLEDEWFKGKYRPEFDE 279

QY 417 -----LPSEEEHCSYVEVTEGEVKNVRLIPSWTIVILVKSMRLKRSFGNPF 463

Db 280 KYDTTLDDVDVAFNDSEHHVTEKKEEPVALNAFELI-----SNSAGNLGNLF 329

QY 464 --EPQARRERSMS--APGNLLVKE 484

Db 330 DSEQEKKRETRFTSKCPKPKETVRK 353

RESULT 12

A56009

serine/threonine-specific protein kinase (EC 2.7.1.-) NPK5 - common tobacco
C;Species: Nicotiana tabacum (common tobacco)

C;Date: 28-Apr-1995 #sequence_revision 28-Apr-1995 #text_change 16-Jun-2000
C;Accession: A56009

R;Muranaka, T.; Banno, H.; Machida, Y.
Mol. Cell. Biol. 14, 2958-2965, 1994

A;Title: Characterization of tobacco protein kinase NPK5, a homolog of Saccharomyces cerevisiae
A;Reference number: A56009; MUID:94217693; PMID:8164654

A;Accession: A56009

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-511 <MUR>

A;Cross-references: GB:D26602; NID:g496384; PIDN:BAA05649.1; PID:g496385
C;Function:

A;Description: catalyzes the formation of peptidyl-serine-phosphate or peptidyl-threonine
C;Superfamily: AMP-activated protein kinase; protein kinase homolog

C;Keywords: ATP; magnesium; phosphotransferase; serine/threonine-specific protein kinase

F;17-271/Domain: protein kinase homology <KIN>

F;25-33/Region: protein kinase ATP-binding motif

F;48,67,142,144/Active site: Lys, Glu, Asp, Lys #status predicted

F;147,151/Binding site: magnesium (Asn, Asp) #status predicted

Query Match 16.6%; Score 436.5; DB 1; Length 511;

Best Local Similarity 33.7%; Pred. No. 1.2e-12;

Matches 96; Conservative 64; Mismatches 94; Indels 31; Gaps 6;

QY 125 LNOYKLOSEIGKAGYGVRLAYNESEDRHYAMKVLKSKLLKQYGFPPRPGSQAAQ 184

Db 16 LRNYKLKGLTGIGSGFKVIAEHTLTGHKVAVKILNRKI----- 55

QY 185 GGPAKQLPLERVYQETAILKLDHVNVLKLEVLDDPAEDNLVLFDLRKGPMVEYPCD 244

Db 56 ---KNMEMEKVREIKLRLFMHPIIRLYEVETPSD--IYVMYIVKSGELFDYIVE 110

QY 245 K-PFSEEQARLYLRDVLGLEYLHCQKIVHRDIKPSNLLLGDDGHVKIADFGVSNQFEGN 303

Db 111 KGRLOEARKPFQIISGVEYCHRNVMVHRDLKPENLLDSKNWVKIADFGLSNIMRDG 170

QY 304 DAQLSSTAGTAPFAPWAPAI SDSQSGSKALDVWATGVTYCYFYGKCPIDDFILALHR 363

Db 171 HF-LKTS CGSPNAAPEVI--SGKYAGPEVDVWSCGVILYALLCGTLFPDENIPNLFK 227

QY 364 KIKNEPVVPFEEPEISEELKDLILKMLDKNPETRIGVDPDKLHPW 408

Db 228 KIKGMISLPS--HLSAGARDLIPRLIIVDPKRMRTIPELRHPW 270

RESULT 13

Tl4822

probable serine/threonine protein kinase (EC 2.7.1.-) SNFL3 - sorghum

N;Alternate names: SNF1 protein kinase homolog SNFL3

C;Species: Sorghum bicolor (sorghum)

C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 21-Jan-2000
C;Accession: Tl4822

R;Annen, F.; Stockhaus, J.

Bot. Acta 111, 137-142, 1998

A;Title: SNFL3 a protein kinase homolog of Sorghum bicolor with a high similarity to the

A;Reference number: Z18217

A;Accession: Tl4822

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: mRNA

A;Residues: 1-461 <ANN>

A;Cross-references: EMBL:Y14274; NID:el301468; PID:el301469

A;Experimental source: cultivar TX 430

C;Genetics:

A;Note: SNFL3

C;Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolog

C;Keywords: phosphotransferase; protein kinase

F;10-266/Domain: protein kinase homology <KIN>

Query Match 16.5%; Score 434; DB 2; Length 461;

Best Local Similarity 29.4%; Pred. No. 1.4e-12;

Matches 111; Conservative 73; Mismatches 135; Indels 58; Gaps 11;

QY 118 DAEDCVOLNOYKLOSBIGKAGYGVRLAYNESEDRHYAMKVLKSKLLKQYGFPPRPPR 177

Db 2 DERRTILMDRYEIGQLGQNFQAKVYFARNLIDGQSVAKMKIDKDKITR-VGL----- 53

QY 178 GSQAAGGPAKQLPLERVYQETAILKLDHVNVLKLEVLDDPAEDNLVLFDLRKG 237

Db 54 -----IVQIKREISIMRLVKHPNLQLFEVM--ASKSKIYFVLEYAKGGE 96

QY 238 VMEVPCDKPFSEEQARLYLRDVLGLEYLHCQKIVHRDIKPSNLLLGDDGHVKIADFGVS 297

Db 97 LFNKISKGKFSBEDVARRYFHLISAVDYCHSRGVVHRDLKPENLLDENENLKVSDFGLS 156

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 12, 2004, 01:20:58 ; Search time 53 Seconds
(without alignments)
496.140 Million cell updates/sec

Title: US-10-690-617-2
Perfect score: 2634
Sequence: 1 MEGGPAVCCQDPAELVERV.....FGEGGKSPFLPGVQDEAAS 505

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues
Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	620	23.5	652	1 SP1 SCHPO	P50526 schizosacch
2	566	21.5	1142	1 PAK1 YEAST	P38990 saccharomyc
3	507.5	19.3	560	1 KGS9 YEAST	P43637 saccharomyc
4	437	16.6	620	1 SNF1 CANAL	P52497 candida alb
5	432	16.4	619	1 SNF1 CANTR	O94168 candida tro
6	426.5	16.2	535	1 K110 ARATH	Q38997 arabidopsis
7	412.5	15.7	756	1 DCK1 MOUSE	Q91lm8 mus musculu
8	410	15.6	408	1 ST6L XENLA	Q91819 xenopus lae
9	407	15.5	611	1 SNF1 CANGA	Q00372 candida gla
10	403.5	15.3	633	1 SNF1 YEAST	P06782 saccharomyc
11	403	15.3	576	1 SNF1 SCHPO	O74536 schizosacch
12	403	15.3	1142	1 GIN4 YEAST	Q12263 saccharomyc
13	402.5	15.3	433	1 DCK1 RAT	O08875 rattus norv
14	401.5	15.2	740	1 DCK1 HUMAN	O15075 homo sapien
15	397	15.1	432	1 STKB XENLA	Q91604 xenopus lae
16	396	15.0	295	1 KMLC DICDI	P25323 dictyosteli
17	394	15.0	648	1 KAPC DICDI	P34099 dictyosteli
18	393	14.9	472	1 KCGG HUMAN	Q13555 homo sapien
19	390.5	14.8	407	1 STK6 XENLA	Q91820 xenopus lae
20	389.5	14.8	552	1 AA2K RAT	Q09137 rattus norv
21	389	14.8	527	1 KCCG RAT	P11730 rattus norv
22	387.5	14.7	552	1 AA2K HUMAN	P54646 homo sapien
23	387	14.7	661	1 ARK5 HUMAN	O60285 homo sapien
24	387	14.7	794	1 K111 HUMAN	Q8tdc3 homo sapien
25	386	14.7	476	1 LOK DROME	O61267 drosophila
26	385.5	14.6	529	1 KCCG MOUSE	Q923c9 mus musculu
27	384.5	14.6	533	1 KCCD RAT	P15791 rattus norv
28	384.5	14.6	1037	1 KCC4 YEAST	P25389 saccharomyc
29	384	14.6	478	1 KCCA HUMAN	Q9ucm7 homo sapien
30	384	14.6	543	1 CHK2 HUMAN	O96017 homo sapien
31	383.5	14.6	478	1 KCCA RAT	P11275 rattus norv
32	383.5	14.5	433	1 STKB HUMAN	Q15831 homo sapien
33	383	14.5	774	1 MRK2 MOUSE	Q05512 mus musculu

34	382.5	14.5	478	1 KCCA MOUSE	P11798 mus musculu
35	382.5	14.5	499	1 KCCD HUMAN	Q13557 homo sapien
36	382	14.5	460	1 CDS1 SCHPO	Q09170 schizosacch
37	378.5	14.4	664	1 KCCB HUMAN	Q13554 homo sapien
38	378	14.4	404	1 KAPC CABEL	P21137 caenorhabdi
39	378	14.4	542	1 KCCB MOUSE	P28652 mus musculu
40	377	14.3	350	1 KAPA MOUSE	P05132 mus musculu
41	377	14.3	481	1 AKT2 HUMAN	P31751 homo sapien
42	375.5	14.3	350	1 KAPA HUMAN	P17612 homo sapien
43	375.5	14.3	550	1 AAK1 HUMAN	Q13131 homo sapien
44	375.5	14.3	623	1 NRC2 NEUCR	O42626 neuospora
45	375.5	14.3	891	1 KINI SCHPO	P22987 schizosacch

ALIGNMENTS

RESULT 1
SSP1 SCHPO
ID SSP1 SCHPO STANDARD; PRT; 652 AA.
AC P50526;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Serine/threonine-protein kinase sspl (EC 2.7.1.-).
GN SSP1 OR SPCC297.03.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95354651; PubMed=7628434;
RA Matsuoka T., Hirata D., Yanagida M., Toda T.;
RT "A novel protein kinase gene sspl+ is required for alteration of
RT growth polarity and actin localization in fission yeast.";
RL EMBO J. 14:3325-3338(1995).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=972;
MEDLINE=21848401; PubMed=11859360;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volkart G., Aert R., Robben J., Grymonprez B.,
RA Weltjens I., Vanstraelen E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Leirach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA Goffeau A., Cadiou E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gallard C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerrutini L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpakowski G.V., Useery D., Barrell B.G., Nurse P.;
RL "The genome sequence of Schizosaccharomyces pombe.";
CC -!- FUNCTION: Involved in actin localization and thus in polarized
CC cell growth.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC -!- SIMILARITY: Belongs to the Ser/Thr family of protein kinases.


```

Best Local Similarity 31.6%; Pred. No. 3e-29;
Matches 151; Conservative 82; Mismatches 135; Indels 110; Gaps 14;

Qy 98 HISPRWRPTTIESHHVAISDAEDCVQ-----LNQYKLOSEIGKGAYG 140
Db 86 HISSSLAKPTTTSFCSSGSKNKKVETNRISITYDPVSRKVLNITYEIKELGHGQHG 145
Qy 141 VVRLAYNESEDRHYAMKVLK--KKLLKQYGFPRPPRGSQAAGGPAKQLLRLVRYQ 198
Db 146 KVKLARDILSKQLVAIKVIDRHEKKQKFFFIK-----SSKISENDKIKR 191
Qy 199 EIALKLLDHNVNKLIVLDDPAEDNLYIVDLLRKGPMVEVPCD-----KPFSE 249
Db 192 EIALMKCHHHVVQVLEVDLRSKTYLVLEYCSRGEVKWCPDCMESDAKGFSLLSF 251
Qy 250 EQARLYLRDVLGLYELHCHQKIVHRDIKPSNLLGDDGHVKIADPGVS-----NQFEGN 303
Db 252 QETREILRGVVLGLYELHYQGIHHRDIKPNALLISGDTGVKISDFGVSLAASSTNSDSS 311
Qy 304 DA----QLSSTAGTAPFAPE--AISDSG-----QSFSGKA-----LDVWATGVTLYC 345
Db 312 ESLDELELAKTVGTPAFAPPEMCLGEDAFTRYNLTKENLFRGSCISFMDIIVAVGVTLYC 371
Qy 346 FVYGKCPIDDFILALHKKIKNEPVVFPPEBEIS-----BELKOLILKMLDK 392
Db 372 LLFGMLPPFSDFLKLFKFIKVIKNDPLKFTPTKEIQSNKVKVSCBEEYEMAKDLLKLEK 431
Qy 393 NPETRIGVDPDKLHPWVT-----KNGEELPS-----EEHCSVVEVTEGEV 434
Db 432 NPQKMTIPALKKPPVSWDFHVPENDEKLLSSVLEQKLRPQCNTQFPFISIKHEL 491
Qy 435 KNSVRLIPSWTIVLVKSM-----LRKRSFGNPFEPQAREERSMSAPGNLLVKEG 485
Db 492 KNAVSGVKIKESVLKSIPLKDPDLGNKNVLYHPTETTRGRGD-----ANVIVSEG 543

RESULT 3
KGS9 YEAST
ID KGS9 YEAST STANDARD; PRT; 560 AA.
AC P43637;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Probable serine/threonine-protein kinase YGL179C (EC 2.7.1.1.-).
GN YGL179C OR G1618.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288c / FY1679;
RX MEDLINE=95397594; PubMed=7668046;
RA Coglievina M., Bertani I., Klina R., Zaccaria P., Bruschi C.V.;
RT "The DNA sequence of a 7941 bp fragment of the left arm of chromosome
RT VII of Saccharomyces cerevisiae contains four open reading frames
RT including the multicopy suppressor gene of the pop2 mutation and a
RT putative serine/threonine protein kinase gene.";
RL Yeast 11:767-774 (1995).
CC -!- SIMILARITY: Belongs to the Ser/Thr family of protein kinases.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X83690; CAA58659.1; -
CC EMBL; Z72701; CAA96891.1; -
CC FIR; S57252; S57252.
CC HSSP; Q00534; 1B18.

```

```

GermOnline; 141227; -.
DR SGD; S0003147; TOS3.
DR InterPro; IPR000719; Prot kinase.
DR InterPro; IPR008271; Ser Thr pkin AS.
DR InterPro; IPR002290; Ser_thr_kinase.
DR Pfam; PF00069; pkinase, 1.
DR ProDom; PD000001; prot_kinase; 1.
DR SMART; SM00220; S_TKC; 1.
DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
DR PROSITE; PS00108; PROTEIN KINASE ST; 1.
DR PROSITE; PS00111; PROTEIN KINASE DOM; 1.
KW Hypothetical protein; Transferase; Serine/threonine-protein kinase;
FT ATP-binding.
KW DOMAIN 50 344 PROTEIN KINASE.
FT NP BIND 56 64 ATP (BY SIMILARITY).
FT BINDING 79 79 ATP (BY SIMILARITY).
FT ACT SITE 189 189 BY SIMILARITY.
FT ACT SITE 189 189 BY SIMILARITY.
SQ SEQUENCE 560 AA; 62090 MW; EPPA0C5ED58B5466 CRC64;

Query Match 19.3%; Score 507.5; DB 1; Length 560;
Best Local Similarity 35.2%; Pred. No. 7.8e-26;
Matches 128; Conservative 58; Mismatches 115; Indels 63; Gaps 10;

Qy 125 LNQYKLOSEIGKGAYGVVRLAYNESEDRHYAMKVLKSKLLKQYGFPRPPRGSQAAGG 184
Db 47 LNNFEILATLNGQYGVKVLARDLGTGALVAIKILNR-----FEKR-----S 88
Qy 185 GPAKQL-LPLSERVQETAILKKLDHNVNKKLIEVLDDPAEDNLYIVFOLLRRGPV----- 238
Db 89 GYSIQLKVENFRVQNEIEVMKRCHEENVVELYELINDPESTKVLVLEYCSRGVPWKCE 148
Qy 239 --MEVPCDKP--FSEEQARLYLRDVLGLYELHCHQKIVHRDIKPSNLLGDDGHVKIADP 294
Db 149 NKMEIKAVGPSILTFQOSRKVVLVDVWSGLEYLHSGQITHRDIKPSNLLISSNGTVKISDF 208
Qy 295 GV-----SNQFSGNDAQL--SSTAGTAPFAPEAISDSGQSFSGKALDVWATGVTLYC 345
Db 209 GVAMSTATGSTNIQSSEHQLKSRALGPPAFAPFELCSTEKEYSCSSAIDVLSIGVTIYC 268
Qy 346 FVYGKCPIDDFILALHKKIKNEPVVFPPEBEISE-----ELKOLILKMLDKN 393
Db 269 LLFGKLPFNANSGLFLPDSILINKPLEFSEYEMLNGATSGITMEYTDKALLKKLLQKD 328
Qy 394 PETRIGVDPDKLHPWVTNKGEEPLPS-----EEEH-----CSVVEVTEGEVKNVR 439
Db 329 PKRIKLADIKVHPFPMCHYGKSDAASVLTNLETHELKVPSPSCRKVELVSLPVNSSF 388
Qy 440 LIPS 443
Db 389 SLDS 392

RESULT 4
SNF1 CANAL
ID SNF1 CANAL STANDARD; PRT; 620 AA.
AC P52497; Q000309;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Carbon catabolite derepressing protein kinase (EC 2.7.1.1.-).
DE SNF1.
GN Candida albicans (Yeast).
OS Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; mitosporic Saccharomycetales; Candida.
OX NCBI_TaxID=5476;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 32354 / B-311;
RX MEDLINE=98053924; PubMed=9393775;
RA Petter R., Chang Y.C., Kwon-Chung K.J.;
RT "A gene homologous to Saccharomyces cerevisiae SNF1 appears to be
RT essential for the viability of Candida albicans.";
RL Infect. Immun. 65:4909-4917 (1997).

```

```

RN  RP  SEQUENCE OF 7-620 FROM N.A.
RC  STRAIN-ATCC 32354 / B-311;
RA  Pectet R., Kwon-Chung K.J.;
RL  Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.
CC  -!- FUNCTION: Essential for release from glucose repression. It
CC  interacts and has functional relationship to the regulatory
CC  protein SNF4. Could phosphorylates CAT8 (By similarity).
CC  -!- SUBCELLULAR LOCATION: Associated with the nuclear membrane (By
CC  similarity).
CC  -!- SIMILARITY: Belongs to the Ser/Thr family of protein kinases. SNF1
CC  subfamily.
-----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
-----
CC  EMBL; L78129; ABA48643.1; -.
CC  EMBL; L39263; AAA92456.1; -.
CC  HSP; Q63450; IA06.
CC  InterPro: IPR000719; Prot kinase.
CC  InterPro: IPR008271; Ser Thr pkin AS.
CC  InterPro: IPR002290; Ser Thr pkinase.
CC  Pfam; PF00069; pkinase; 1.
CC  ProDom; PD000001; Prot kinase; 1.
CC  SMART; SM00220; S_TKC; 1.
CC  PROSITE; PS00107; PROTEIN KINASE ATP; 1.
CC  PROSITE; PS00108; PROTEIN KINASE ST; 1.
CC  PROSITE; PS50011; PROTEIN KINASE DOM; 1.
CC  Transferrase; Serine/threonine-protein kinase; ATP-binding;
CC  Phosphorylation; Carbohydrate metabolism; Nuclear protein.
CC  DOMAIN 16 29 POLY-HIS.
CC  FT  DOMAIN 53 305 PROTEIN KINASE.
CC  FT  NP BIND 59 67 ATP (BY SIMILARITY).
CC  FT  BINDING 82 82 ATP (BY SIMILARITY).
CC  FT  ACT SITE 175 175 BY SIMILARITY.
CC  FT  MOD_RES 208 208 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
CC  FT  CONFLICT 7 8 PQ -> AR (IN REF. 2).
CC  FT  CONFLICT 228 234 AGEVDV -> SSVQKLM (IN REF. 2).
CC  FT  CONFLICT 242 255 YVMCGRLPFDEF -> GMSCCVVDVHMTSS (IN
CC  REF. 2).
CC  FT  CONFLICT 387 387 A -> R (IN REF. 2).
CC  FT  CONFLICT 416 416 R -> A (IN REF. 2).
CC  FT  CONFLICT 494 494 S -> L (IN REF. 2).
CC  SQ  SEQUENCE 620 AA; 70005 MW; 1806C652B5061D2B CRC64;

Query Match 16.6%; Score 437; DB 1; Length 620;
Best Local Similarity 29.1%; Pred. No. 3.5e-21;
Matches 122; Conservative 78; Mismatches 137; Indels 82; Gaps 12;

QY 124 QLNQYKQSEIGKYGAVVRLAYNESEDRHYAMKVLKSKKLLKQYGFPRPPRGSAQAQ 183
D 49 RIGRYQILKTLGSGFGKVLKQHLGTGQVALKILNRKTLAK-----SDMQ 95
QY 184 GGPAAQLPLRYVYQETAILKLDHNVVVKLIEVLDPAEDNLYLVFDLLRKGVMVPC 243
D 96 G-----RVERISYLRLLRHPIIKLYDVI--KSKDEIMVIEFAGKELFDYIVQ 143
QY 244 DKPFSEQARLYLRDVLIGLEYLHCQKIVHRDKPSNLLIGDGDGHVKIADFGVSN-QFEG 302
D 144 RGMPEDEARRFFQIIAAVEYCHRHKIVHRDLKPENLLDDQNLQNLKIDFGLSNITMDG 203
QY 303 NDQLQSTAGTAPM-APEAISDGSQSFSGKALDWATGVTLYCFYVKGCPFDIFDAL 361
D 204 N--FLKTSQSPNMPAPEVI--SGKLYAGFVDVMSAGVILYVLMCGRLPFDFEPAL 259
QY 362 HRKIKPEVPVPEEPISEBLKLLKMLDKNPETRIGV-----PDIKLHPW 408
D 260 FKXISNGVYTLN--YLSAGAKHLLTRMLVNVNPLNRITHEIMEDDWFQDMDFYLLPPD 317

```

```

QY 409 VTKNGEPLPSEBEHCSEVVEVTEGEVKNVRLIPSWTTLVLVKMLKRSGFNPEPOAR 468
D 318 LSKNKNKSIDVDENVIRALSVTMGYDRDC-----KIVNVIEKANKQVAGNSSQSK 370
QY 469 REERSMSPAGNLLVKEG-----FGSGKSPPELPVQOE 500
D 371 SSNEILDA--YLLMKENHALVKDLKSKSENIESFLSQSPPPSPFNRGSTSSAPGVQ 427

RESULT 5
SNF1_CANTR STANDARD; PRT; 619 AA.
AC Q94168;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Carbon catabolite derepressing protein kinase (EC 2.7.1.-).
GN SNF1.
OS Candida tropicalis (Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; mitosporic Saccharomycetales; Candida.
OX NCBI_TaxID=5482;
RN [1]
RP SEQUENCE FROM N.A.
RA Kanai T., Ogawa K., Ueda M., Tanaka A.;
RT 'Genetic evaluation of the function of SNF1 in Candida tropicalis.';
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Essential for release from glucose repression. It
CC interacts and has functional relationship to the regulatory
CC protein SNF4. Could phosphorylates CAT8 (By similarity).
CC -!- SUBCELLULAR LOCATION: Associated with the nuclear membrane (By
CC similarity).
CC -!- SIMILARITY: Belongs to the Ser/Thr family of protein kinases. SNF1
CC subfamily.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL; AB024535; BAA75889.1; -.
CC HSP; Q63450; IA06.
CC InterPro: IPR000719; Prot kinase.
CC InterPro: IPR008271; Ser Thr pkin AS.
CC InterPro: IPR002290; Ser Thr pkinase.
CC Pfam; PF00069; pkinase; 1.
CC ProDom; PD000001; Prot kinase; 1.
CC SMART; SM00220; S_TKC; 1.
CC PROSITE; PS00107; PROTEIN KINASE ATP; 1.
CC PROSITE; PS00108; PROTEIN KINASE ST; 1.
CC PROSITE; PS50011; PROTEIN KINASE DOM; 1.
CC Transferrase; Serine/threonine-protein kinase; ATP-binding;
CC Phosphorylation; Carbohydrate metabolism; Nuclear protein.
CC DOMAIN 17 23 POLY-HIS.
CC FT  DOMAIN 52 303 PROTEIN KINASE.
CC FT  NP BIND 58 66 ATP (BY SIMILARITY).
CC FT  BINDING 81 81 ATP (BY SIMILARITY).
CC FT  ACT SITE 174 174 BY SIMILARITY.
CC FT  MOD_RES 207 207 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
CC SQ  SEQUENCE 619 AA; 70323 MW; 0FC1FC3DCB706D7 CRC64;

Query Match 16.4%; Score 432; DB 1; Length 619;
Best Local Similarity 28.4%; Pred. No. 7.3e-21;
Matches 119; Conservative 80; Mismatches 138; Indels 82; Gaps 11;

QY 124 QLNQYKQSEIGKYGAVVRLAYNESEDRHYAMKVLKSKKLLKQYGFPRPPRGSAQAQ 183
D 48 RIGRYQILKTLGSGFGKVLKQHLGTGQVALKILNRKTLAK-----SDMQ 94

```



```

DR InterPro; IPR001245; Tyr_kinase.
DR InterPro; IPR000449; UBA_domain.
DR Pfam; PF02149; Kx1; 1.
DR Pfam; PF00069; pkinase; 1.
DR Pfam; PF00627; UBA; 1.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TKc; 1.
DR SMART; SM0165; UBA; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE; PS0030; UBA; 1.
DR Transferase; Serine/threonine-protein kinase; ATP-binding;
KW Phosphorylation; Alternative splicing.
FT DOMAIN 42 294
FT NP_BIND 48 56
FT FT 315 355
FT ACT_SITE 165 165
FT BINDING 71 71
FT MOD_RES 198 198
FT VARSPIC 1 23
FT SEQUENCE 535 AA; 61181 MW; PFFC383223FD8317 CRC64;

Query Match 16.2%; Score 426.5; DB 1; Length 535;
Best Local Similarity 34.0%; Pred. No. 1.4e-20;
Matches 97; Conservative 59; Mismatches 98; Indels 31; Gaps 6;

QY 125 LNQYKLGSEITGKGYGVRLAYNESEDRHYAMKVLKSKLLKQYGFPRPPRGSAQAQ 184
Db 39 LPNYKLGRTLGSGFGRVKAIEHALTGHKVAIKLNRKI----- 78

QY 185 GPAQLPLSRVQVEIALLKLDHVNVLVLEVLDDPAEDNLYLVFLLRKGPVMEVPCD 244
Db 79 ---KNMEMEKVRREIKILFLPHILRIYEVETPD--IYLWVEYVNSGELFDYIVE 133

QY 245 K-PFSEQARLYLRDVLGLEYLHCQKIVHRDIKPSNLLGDDGHVKIADFGVSNQPEGN 303
Db 134 KGRLQDEARNFQQIISGVYCHRNVHRDLKPNLLDLSKCNKIADFGSLNMRDG 193

QY 304 DAQLSSTAGTAPAFMAPAISDSQSGSKALDWATGVTLVYCFYVYKCPFDIFILALHR 363
Db 194 HF-LKTSQSPNVAAPPEVI--SGKLYAGPEVDVMSGCVILYALLCGTLPEDDENIPNLFK 250

QY 364 KIKNEPVVPEEPPEISEELKDLKMLDKNPETRIGVDPDKLHPW 408
Db 251 KIKGIYTLFS--HLSPGARDLIPRLVLDVPMKRVITPEIRQHPW 293

RESULT 7
ID DCK1_MOUSE STANDARD; PRT; 756 AA.
AC Q9ULM8;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Serine/threonine-protein kinase DCAMK1 (EC 2.7.1.-) (Doublecortin-
DE like and CAM kinase-like 1).
DE DCAMK1 OR DCLK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=20004649; PubMed=10533048;
RA Burgess H.A., Martinec S., Reiner O.;
RT "KIAA0369, doublecortin-like kinase, is expressed during brain
RT development.";
RL J. Neurosci. Res. 58:567-575 (1999).
-|- FUNCTION: PROBABLE KINASE THAT MAY BE INVOLVED IN A CALCIUM-

```

```

CC SIGNALING PATHWAY CONTROLLING NEURONAL MIGRATION IN THE DEVELOPING
CC BRAIN. MAY ALSO PARTICIPATE IN FUNCTIONS OF THE MATURE NERVOUS
CC SYSTEM (by similarity).
CC -|- SIMILARITY: Belongs to the Ser/Thr family of protein kinases. CamK
CC subfamily.
CC -|- SIMILARITY: Contains 2 doublecortin domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF155819; AAF26673.1; -.
CC HSPP; Q63450; IA06.
CC MGD; MGI:1330861; Dcamk1l.
CC InterPro; IPR003533; DCX.
CC InterPro; IPR000719; Prot_kinase.
CC InterPro; IPR008271; Ser_Ehr_pkin_AS.
CC InterPro; IPR002290; Ser_thr_pkinase.
CC Pfam; PF03607; DCX; 2.
CC Pfam; PF00069; pkinase; 1.
CC ProDom; PD000001; Prot_kinase; 1.
CC SMART; SM00537; DCX; 2.
CC SMART; SM00220; S_TKc; 1.
CC PROSITE; PS00309; DC; 2.
CC PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
CC PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
CC PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
CC Transferase; Serine/threonine-protein kinase; ATP-binding; Repeat;
CC Neurogenesis. 57 143
FT DOMAIN 186 269
FT DOMAIN 298 358
FT DOMAIN 406 663
FT NP_BIND 412 420
FT BINDING 435 435
FT ACT_SITE 527 527
FT SEQUENCE 756 AA; 84153 MW; 3D1DBF18C23129F2 CRC64;

Query Match 15.7%; Score 412.5; DB 1; Length 756;
Best Local Similarity 27.6%; Pred. No. 1.8e-19;
Matches 140; Conservative 71; Mismatches 199; Indels 97; Gaps 20;

QY 35 PEPTNRGVDP-----PPRARAASVTPGSTSRLLPAPSL-SARKUSLQE 77
Db 299 PGPSRRSKSPASTSVNGTPTGSQLSTPRSGKSPSPSTPSGLRKQRIHQSGSSTLS 358

QY 78 RPAGSYLEAQAQPYATGPAS-----HISPRWRPRTIESHHVAISDAEDCVL-----NQ 127
Db 359 TKVCSSMDNDGDP---GEGDELGRHSLQGRMR-----ESEEQFIPATITER 405

QY 128 YKLQSEIGKGYGVRLAYNESEDRHYAMKVLKSKLLKQYGFPRPPRGSAQAQCPA 187
Db 406 YKVGRTIGDGNFAVVEKCIETAREYALKIKKSK-----RGKE----- 446

QY 188 KQLPLERVQVEITAILKLDHVNVLKIEVLDDPAEDNLYLVFLLRKGPVME-VPCDKP 246
Db 447 -----HMIQNEVSILRRVKHPNIVILLIEMDVPTD--LYLVNELVKGGLDFAITSTSK 498

QY 247 FSEQEARLYLRDVLGLEYLHCQKIVHRDIKPSNLLGDP--DG--HVK1ADFGVSNQFEG 302
Db 499 YTERDASGLMYNLASAIAKVLHSLNIVHRDIKENLLEVEHQDQSKLKGDFGLATIVDG 558

QY 303 NDAQLSSTAGTAPAFMAPAISDSQSGSKALDWATGVTLVYCFYVYKCPFP--IDDFILA 360
Db 559 ---PLYTVCGTPTVYVPEIETGY---GLKWDIAAGVITYILLCGFPFRSGDDQEV 612

QY 361 LHRKIKNEPVVPEE--PISIELKDLILKMLDKNPETRIGVDPDKLHPWTKNGEPLP 418
Db 613 LFDQILMGQVDFPSPFYWDNVSASAKELINMILLVNVVDQRFSAVQVLEHPWVNDG---LP 669

```

```

QY 419 SEEBCHSVVTEGEVKNVSLIPSWTTVLVKMLKRKSGNPFEPQARREMSAPG 478
Db 670 ENEHOLSVA---GKIKKHFTGPKPSSTAAGSVIATTALDK--ERQVFRRRNQD--- 720
QY 479 NLLYKEGEGGKSPELPGVQDEAAS 505
Db 721 ---VRSRYKAOPAPPELNSSESDYSPS 744

RESULT 8
ST6L_XENLA
ID ST6L_XENLA STANDARD; PRT; 408 AA.
AC Q91819;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Serine/threonine protein kinase Eg2-like (EC 2.7.1.37) (p46X1Eg22).
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Eupidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Egg;
RA Roghi C., Le Guellec R., Paris J., Couturier A., Philippe M.;
RT "Eg2, selected by differential screening encodes a new Xenopus protein
RT kinase family.";
RL Submitted (OCT-1992) to the EMBL/GenBank/DBJ databases.
CC -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC -!- SUBCELLULAR LOCATION: Localizes to the spindle pole during mitosis
CC especially from prophase through anaphase. Partially colocalised
CC with gamma tubulin in the centrosome, from S to M phase (By
CC similarity).
CC -!- PTM: Phosphorylated (By similarity).
CC -!- SIMILARITY: Belongs to the Ser/Thr family of protein kinases.
CC Aurora subfamily.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).

DR EMBL; Z17206; CAA78914.1; ALT_INIT.
DR HSSP; P24941; 1AQL.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR008271; Ser_thr_pkin_AS.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TKc; 1.
DR SMART; SM00219; TyrKc; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW Cell cycle; Transferase; Serine/threonine-protein kinase; ATP-binding;
Phosphorylation.
FT DOMAIN 140 390 PROTEIN_KINASE.
FT NP_BIND 146 154 ATP (BY SIMILARITY).
FT BINDING 169 169 ATP (BY SIMILARITY).
FT ACT_SITE 263 263 BY SIMILARITY.
SQ SEQUENCE 408 AA; 46461 MW; 97FeA69C7357AE8 CRC64;

Query Match 15.6%; Score 410; DB 1; Length 408;
Best Local Similarity 29.9%; Pred. No. 1.2e-19;
Matches 126; Conservative 61; Mismatches 155; Indels 80; Gaps 16;

37 PFRNGVDPPPARAASVTPG-STSRLL-----PARPSLSARK--LSLQRPAGSVLEAQA 88
QY

```

```

Db 24 PKRIPVSQPPSTQVRPPVTGVSQAQILGPSNVPPQRMQAOKPVLNSOKPTAQQLLR--- 80
QY 89 GPYATGPASH-----ISPRAWRRPTTESHH-----VAISDAED----- 121
Db 81 -----PATHGHQTSKPOGPNENRNPOQTSSTSTPMKKKSGTDQCKTILAVPEKGGKKQ 134
QY 122 -CVQLNQYKLSQSEIGKYGAYVVRLAYNESEDRHYAMKVLKKLLKQYGFRRPPRRSQ 180
Db 135 WC--LEDFEIGRPLGKGFKNVYLARESERESKFILALKVLFK-----SQ 175
QY 181 AAQGGPAKOLLPLERVYQEIALLKLDHVVNVKLVLELDLPAEDNLVLFVDFLLRKGPM- 239
Db 176 LEKAGVEHQL-----RREEIQSHLRHPNLTILYGYFHDAS--RVYILDYAPGGELFR 227
QY 240 EVPCKPFSEQARLYLRDVLILGLEYLHCQKTVHRDIPKPSNLLIGDDGHVKTADFGVSNQ 299
Db 228 ELQKTRFDORSALYIKQLAEALLYCHSKVIHRDIPENLLLSNGSELKIADEGWS-- 285
QY 300 FEGNDAQLSSTAGTPAFMAPEAISDSGQSFSGKALDWATGVTLVCFYVGKCPFIDDFIL 359
Db 286 VHAPSSRRRTTLCGLDYLPPEMI--EGRMHDEK-VDLMSLGLVLCYEFVLGKPPFETDTHQ 342
QY 360 ALHRKIKNEPVVFPEEPEISEBELKDLILKMLDKNPETRIGVDPDIKLHPVTKNGEPIPS 419
Db 343 ETYRRISK--VEFOYPPYVSEAKDLVSKLLKHNPNHRLPLKGVLEHPWIVKNSQQPKKK 400
QY 420 EE 421
Db 401 DE 402

RESULT 9
SNF1_CANGA
ID SNF1_CANGA STANDARD; PRT; 611 AA.
AC Q00372;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Carbon catabolite derepressing protein kinase (EC 2.7.1.1)-.
GN SNF1.
OS Candida glabrata (Yeast) (Torulopsis glabrata).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; mitosporic Saccharomycetales; Candida.
OX NCBI_TaxID=5478;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NCCLS84;
RX MEDLINE=97101049; PubMed=8945576;
RA Petter R., Kwon-Chung K.J.;
RT "Disruption of the SNF1 gene abolishes trehalose utilization in the
RT pathogenic yeast Candida glabrata.";
RL Infect. Immun. 64:5269-5273(1996).
CC -!- FUNCTION: Essential for release from glucose repression. It
CC interacts and has functional relationship to the regulatory
CC protein SNF4. Could phosphorylates CAR8 (By similarity).
CC -!- SUBCELLULAR LOCATION: Associated with the nuclear membrane (By
CC similarity).
CC -!- SIMILARITY: Belongs to the Ser/Thr family of protein kinases. SNF1
CC subfamily.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).

DR EMBL; L78130; AAB48642.1; -.
DR HSSP; P24941; 1HCL.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR008271; Ser_thr_pkin_AS.

```



```

GIN4_YEAST      STANDARD;      PRT; 1142 AA.
ID  GIN4_YEAST
AC  Q12263;
DT  01-NOV-1997 (Rel. 35, Created)
DT  01-NOV-1997 (Rel. 35, Last sequence update)
DT  30-MAY-2000 (Rel. 39, Last annotation update)
DE  Serine/threonine-protein kinase GIM4 (EC 2.7.1.-).
GN  GIN4 OR YDR507C OR D9719.13.
OS  Saccharomyces cerevisiae (Baker's yeast).
OC  Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC  Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX  NCBI_TaxID=4932;
RN  [1]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=99030835; PubMed=9813093;
RA  Longline M.S., Fares H., Pringle J.R.;
RT  "role of the yeast Gln4p protein kinase in septin assembly and the
RT  relationship between septin assembly and septin function.";
RL  J. Cell Biol. 143:719-736 (1998).
[2]
RP  SEQUENCE FROM N.A.
RA  Dietrich F.S., Mulligan J., Allen E., Araujo R., Aviles E., Berno A.,
RA  Carpenter J., Chen E., Cherry J.M., Chung E., Duncan M.,
RA  Hunicke-Smith S., Hyman R., Komp C., Lashkari D., Lew H., Lin D.,
RA  Mosedale D., Nakahara K., Nanath A., Oefner P., Oh C., Petel F.X.,
RA  Roberts D., Schramm S., Schroeder M., Shogren T., Shroff N.,
RA  Winant A., Yelton M., Botstein D., Davis R.W.;
RL  Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.
CC  -!- FUNCTION: May play a role in septin assembly.
CC  -!- SIMILARITY: Belongs to the Ser/Thr family of protein kinases. NIM1
CC  subfamily.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
CC  EMBL; U33140; AAA75513.1; -.
CC  DR  EMBL; U33057; AAB64949.1; -.
CC  DR  PIR; S59359; S59359.
CC  DR  HSP; Q63450; 1A06.
CC  DR  Germonline; 140999; -.
CC  DR  SGD; S0002915; GIN4.
CC  DR  GO; GO:0005935; C:bud neck; IDA.
CC  DR  GO; GO:0007117; P:bud growth; IGI.
CC  DR  GO; GO:0001335; P:septin checkpoint; IGI.
CC  DR  GO; GO:0000921; P:septin ring assembly; IGI.
CC  DR  InterPro; IPR000719; Prot_kinase.
CC  DR  InterPro; IPR008271; Ser_thr_pkin_AS.
CC  DR  InterPro; IPR002290; Ser_thr_pkinase.
CC  DR  Pfam; PF00069; pkinase; 1.
CC  DR  PRINTS; PR00109; TYRKINASE.
CC  DR  ProDom; PD000001; Prot_kinase; 1.
CC  DR  SMART; SMO0220; S_TKC; 1.
CC  DR  PROSITE; PS00107; PROTEIN KINASE ATP; 1.
CC  DR  PROSITE; PS00108; PROTEIN KINASE ST; 1.
CC  DR  PROSITE; PS00011; PROTEIN KINASE DOM; 1.
CC  KW  Transferase; Serine/threonine-protein kinase; ATP-binding.
CC  FT  DOMAIN 19 289 PROTEIN KINASE.
CC  FT  NP_BIND 25 33 ATP (BY SIMILARITY).
CC  FT  BINDING 48 48 ATP (BY SIMILARITY).
CC  FT  ACT_SITE 156 156 BY SIMILARITY.
CC  SQ  SEQUENCE 1142 AA; 129857 MW; EC16FF4BB49DD811 CRC64;
Query Match 15.3%; Score 403; DB 1; Length 1142;
Best Local Similarity 32.4%; Pred. No. 1.2e-18;
Matches 119; Conservative 57; Mismatches 143; Indels 48; Gaps 12;
QY 125 LNQYKLOQSEIGKGVGVVRLAYNESEDRHYAMKVLKSKLLKQYGFPRPPRGSQAQG 184

```

```

Db 16 IGFWKLGELGLGSGTKVQLARNGSTGQEAARVSKAVF-----NTGVSGTIVG 67
QY 185 GPAKQLLP--LERYOEIAILKKLDHVNVLKLEVLDDPAEDNLYLVDFLLRKGPMVEVP 242
Db 68 STTPDALPYGIER--EIIIMKLLNHPNVLRYDVWE--TNTDLVLYLVEAERGFENLL 122
QY 243 CDK-PFSEEQARLYLRDVLGLGLEYLHCQKIVHRDIKPSNLLLGDDGHVKTADGVSQFQ 301
Db 123 VERGPLPEHAIRFFQIIGVSYCHALGIVHRDLKPNLLOHKNYKIKIADFGMA-ALE 181
QY 302 GNDALQSSTAGTAPAFMAPRAISDGSFGSKALDVWATGVTLYCFVYVGKQFP--IDDFIL 359
Db 182 TECKLLETSCGSPHYAAPEIV--SGIPYQGFASDVMSGVLFALLTGRLPFDEEDGNIR 239
QY 360 ALHRKIKNEVPVPEPEPEISEELKDLILKMLDKNPETRICVPDIKLPWYTK-----N 412
Db 240 TLLLKVGKGFEMPSDDEISREAQDLIRKILTVDPERRIKTRDILKHPLLQKYPISIRDSK 299
QY 413 GEPELPSEEEHCSVVEVTEGEVKNVRLIPSWTIVL-----VKSMLEKRSFGNPFE 464
Db 300 SIRGLPREDIYLTPL-----SESNSIDATILQNLVILWHGRDPEGIKEKL-R-----E 347
QY 465 PQARREE 471
Db 348 PGANAEX 354
RESULT 13
DCK1_RAT
ID  DCK1_RAT      STANDARD;      PRT; 433 AA.
AC  O08875;
DT  30-MAY-2000 (Rel. 39, Created)
DT  30-MAY-2000 (Rel. 39, Last sequence update)
DT  30-MAY-2000 (Rel. 39, Last annotation update)
DE  Serine/threonine-protein kinase DCAMKL1 (EC 2.7.1.-) (Doublecortin-
DE  like and CAM kinase-like 1) (Calcium/calmodulin-dependent protein
DE  kinase type I-like CPG16).
DE  DCAMKL1 OR CPG16.
GN  Rattus norvegicus (Rat).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX  NCBI_TaxID=101116;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=Wistar;
RX  MEDLINE=98364306; PubMed=9699150;
RA  Hevroni D., Rattner A., Bundman M., Lederfein D., Gabarah A.,
RA  Mangelus M., Silverman M.A., Kedar H., Naor C., Kornuc M., Hanoch T.,
RA  Seger R., Theill L.E., Nedivi E., Richter-Levin G., Citri Y.;
RA  "Hippocampal plasticity involves extensive gene induction and multiple
RA  cellular mechanisms.";
RT  J. Mol. Neurosci. 10:75-98 (1998).
RL  -!- FUNCTION: PROBABLE KINASE THAT MAY BE INVOLVED IN A CALCIUM-
CC  SIGNALING PATHWAY CONTROLLING NEURONAL MIGRATION IN THE DEVELOPING
CC  BRAIN. MAY ALSO PARTICIPATE IN FUNCTIONS OF THE MATURE NERVOUS
CC  SYSTEM (BY SIMILARITY).
CC  -!- SIMILARITY: Belongs to the Ser/Thr family of protein kinases. CamK
CC  subfamily.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
CC  EMBL; U78857; AAC99476.1; -.
CC  DR  HSP; Q63450; 1A06.
CC  DR  InterPro; IPR000719; Prot_kinase.
CC  DR  InterPro; IPR008271; Ser_thr_pkin_AS.
CC  DR  InterPro; IPR002290; Ser_thr_pkinase.

```



```
QY 125 LNOYKLOSEIGKAYGVVRLAYNESEDRHYAMKVL SKKKLLKQYGFPRPPRPSQAAG 184
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
49 VGKYLMDLLGEGSYGKVKEMLDSTLCRRAYKILKKKL-----RRIP-----N 93
QY 185 GPAKQLLPLERVYQEIAILKKLDHVNKKLIEVLDDPAEDNLYLVFDLLRKG--PVMEVP 242
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
94 GEA-----NVKKEIQLLRRLRHRNVQLVDVLYNEEKQKMYMWEYCVCGMQEMLDSV 146
QY 243 CDKPFSEQARLYLRDVLGLYLELHCQKIVHRDIKPSNLLGDDGHVKIADFGVSNQF-- 300
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
147 QDKHPVPFOAHGYFCQLIDGLEYLELHSGQIVHKDIKPGNLLTTDGTCLKISDLGVAELHP 206
QY 301 --EGNDAQLSSTACTPAFWAPEAISDSGQSFSCKALDVWATGVTLVCFVYKCPFIDDFI 358
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
207 FAEQDTCRTSQ--GSPAFQPPPE--IANGLDTSFGFKVDIWSAGVTLYNITTGLYPFEGDNI 263
QY 359 LALHRKIKNEPVVPFPEEPISEELKDLILKMLDKNPETRIQVDPDIKLPWVTKNKEE--- 415
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
264 YKLFENIGKGYSIPE--ECGPLLSDLIRGMLLEYDPAKRFSIQQIRQHNWFRKKHPHMDP 321
QY 416 --PLPSEEE 422
Db : : : : :
322 IVPIPPSPE 330
```

Search completed: July 12, 2004, 04:51:13
Job time : 56 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 12, 2004, 01:17:43 ; Search time 102 Seconds
(without alignments)

1398.887 Million cell updates/sec

Title: US-10-690-617-2

Perfect score: 2834
Sequence: 1 MEGAPAYCCDPAELRV.....FGEKGKSPFLGVQDEFAAS 505

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04.*

- 1: Geneseqp1980s.*
- 2: Geneseqp1990s.*
- 3: Geneseqp2000s.*
- 4: Geneseqp2001s.*
- 5: Geneseqp2002s.*
- 6: Geneseqp2003s.*
- 7: Geneseqp2003bs.*
- 8: Geneseqp2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2634	100.0	505	5	ABB83054 Human kin
2	2626	99.7	505	5	AAB19161 Human kin
3	2626	99.7	505	7	ADE61939 Human Pro
4	2626	99.7	505	7	ADE61943 Human Pro
5	2619	99.4	582	5	ABP69334 Human pol
6	2611	99.1	505	5	ABU65068 Human NOV
7	2595	98.5	513	4	AAU03510 Human pro
8	2595	98.5	513	4	AAE04361 Human kin
9	2591	98.4	561	5	AAU79458 Human nov
10	2585	98.1	543	5	ABU65069 Human NOV
11	2560	97.2	503	5	ABU65067 Human NOV
12	2468	93.7	505	5	ABB83055 Rat amino
13	2468	93.7	505	7	ADE61937 Rat Prote
14	2468	93.7	505	7	ADE61941 Rat Prote
15	1537.5	58.4	540	3	AAV44239 Human cel
16	1525.5	57.9	556	7	ADB75228 Prostate
17	1509	57.3	680	4	AAU40450 Human pol
18	1509	57.3	680	4	AAU40449 Human pol
19	1504	57.1	588	4	AAU38663 Human pol
20	1498	56.9	588	4	AAU38659 Novel pro
21	1472.5	55.9	417	5	ABG61855 Prostate
22	1472.5	55.9	417	7	ADB75230 Prostate
23	1342	50.9	498	7	AUC99093 Human RPP
24	1308.5	49.7	545	4	AAU38664 Human pol
25	923	35.0	210	4	AAU25244 Human pro

26	881	33.4	338	4	ABB66416	Abb66416 Drosophil
27	648	24.6	423	6	ABU11714	Abu11714 Human MDD
28	611	23.2	140	4	AAM25471	Aam25471 Human pro
29	566	21.5	1142	6	ABR52943	ABr52943 Protein s
30	526	20.0	250	5	ABB78798	Abb78798 Serine/th
31	452	17.2	570	6	ABR40710	ABr40710 Zea may9
32	436.5	16.6	511	2	AAR40842	Aar40842 SHPP. 8/2
33	430.5	16.3	523	3	AAB03425	Aab03425 Wheat put
34	430.5	16.3	523	6	ABR40719	ABr40719 Triticum
35	429.5	16.3	422	3	AAB03419	Aab03419 Soybean p
36	429.5	16.3	422	6	ABR40713	ABr40713 Glycine m
37	428.5	16.3	438	3	AAB03421	Aab03421 Soybean p
38	428.5	16.3	438	6	ABR40715	ABr40715 Glycine m
39	428	16.2	441	3	AAB03423	Aab03423 Soybean p
40	428	16.2	441	6	ABR40717	ABr40717 Glycine m
41	427	16.2	254	5	ABB78797	Abb78797 Protein k
42	426.5	16.2	512	6	ABR44015	ABr44015 Human SNF
43	426.5	16.2	1349	4	AAG70854	Aag70854 C albican
44	424.5	16.1	579	3	AAB03417	Aab03417 Corn puta
45	424.5	16.1	579	6	ABR40709	ABr40709 Zea may5

ALIGNMENTS

RESULT 1

ABB83054
ID ABB83054 standard; protein; 505 AA.

AC ABB83054;

DT 27-AUG-2002 (first entry)

DE Human kinase amino acid sequence.

KW Human; kinase protein; calcium/calmodulin-dependant protein kinase; eye;
KW retinoblastoma; brain; kinase modulator.

OS Homo sapiens.

FT Key Location/Qualifiers

FT Region 26..29 /label= casein_kinase_II_phosphorylation_site

FT Region 58..60 /label= protein_kinase_C_phosphorylation_site

FT Region 69..71 /label= protein_kinase_C_phosphorylation_site

FT Region 71..74 /note= "cAMP- and cGMP-dependant protein kinase phosphorylation site"

FT Region 74..77 /label= casein_kinase_II_phosphorylation_site

FT Region 82..85 /label= casein_kinase_II_phosphorylation_site

FT Region 100..102 /label= protein_kinase_C_phosphorylation_site

FT Region 105..108 /note= "cAMP- and cGMP-dependant protein kinase phosphorylation site"

FT Region 117..120 /label= casein_kinase_II_phosphorylation_site

FT Region 134..157 /note= "protein kinase ATP-binding region signature"

FT Region 147..150 /label= N-glycosylation_site

FT Region 160..162 /label= protein_kinase_C_phosphorylation_site

FT Region 178..183 /label= N-myristoylation site

FT Region 271..283 /note= "serine/threonine protein kinase active-site signature"

FT Region 326..331

FT Region /label= N-myristoylation site O
 FT 330..332
 FT /label= protein_kinase_C_phosphorylation_site
 FT Misc-difference 375
 FT /note= "wild-type Glu is replaced by Gly as a result of
 FT an SNP (single nucleotide polymorphism) in the encoding
 FT DNA"
 FT 419..422
 FT /label= casein_kinase_II_phosphorylation_site
 FT 425..428
 FT /label= casein_kinase_II_phosphorylation_site
 FT 430..433
 FT /label= casein_kinase_II_phosphorylation_site
 FT 437..439
 FT /label= protein_kinase_C_phosphorylation_site
 FT 455..458
 FT /note= "cAMP- and cGMP-dependant protein kinase
 FT phosphorylation site"
 FT 485..492
 FT /label= ATP/GTP-binding_site_motif_A
 FT /note= "P-loop"
 FT
 FT WO200224920-A2.
 PN
 PD 28-MAR-2002.
 XX
 XX 19-SEP-2001; 2001WO-US029161.
 PD
 PF WPI; 2002-404955/43.
 DR N-PSDB; ABL58698.
 XX
 XX Novel peptide designated as human kinase useful as target for diagnosing
 PT a disease or predisposition to the disease mediated by the peptide.
 FT
 XX Claim 1a; Fig 2; 89pp; English.
 XX
 CC The invention relates to an isolated peptide designated human kinase
 CC (HK), that has homology to members of the calcium/calmodulin-dependent
 CC protein kinase kinase subfamily. The mechanism of action of the protein
 CC of the invention is that of a kinase modulator. The human kinase of the
 CC invention is useful for creating a pharmaceutical composition for
 CC treating a disease or condition mediated by the human kinase. HK is also
 CC useful to provide a target for diagnosing a disease or predisposition to
 CC disease mediated by HK, and is also useful in pharmacogenic analysis. HK
 CC is useful for treating a disorder characterised by absence of
 CC inappropriate or unwanted expression of HK, also as an immunogen to raise
 CC antibodies by administering HK to a mammalian organism e.g. rat, rabbit
 CC or mouse. Nucleic acids of the invention are useful as hybridisation
 CC probes for mRNA, transcript/cDNA and genomic DNA. Recombinant host cells
 CC expressing a native HK are useful for assaying compounds that stimulate
 CC or inhibit HK function. Nucleic acids of the invention are also useful
 CC for producing transgenic animals. Experimental data indicates that kinase
 CC proteins of the present invention are expressed in humans in the eye
 CC (retinoblastomas) and brain. The current sequence represents the amino
 CC acid sequence of the human kinase of the invention
 XX
 SQ Sequence 505 AA;
 Query Match 100.0%; Score 2634; DB 5; Length 505;
 Best Local Similarity 100.0%; Pred. No. 2.8e-236;
 Matches 505; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 1 MEGGPAVCCDPRAELVERVAALDVTHLEADGGPEPTNGVDPPPRARAASVIPGSTR 60
 1 MEGGPAVCCDPRAELVERVAALDVTHLEADGGPEPTNGVDPPPRARAASVIPGSTR 60

QY 61 LLPAPSLSARKLSLOERPAGSYLEAQAGPYATGPASHISPRARRPTTIESHVAISDAE 120
 DB 61 LLPAPSLSARKLSLOERPAGSYLEAQAGPYATGPASHISPRARRPTTIESHVAISDAE 120
 QY 121 DCVOLNQYKLSQSEIGKAGVVRLAYNESDRHYAMKVLSSKKLLKQYGFPRPPRGSQ 180
 DB 121 DCVOLNQYKLSQSEIGKAGVVRLAYNESDRHYAMKVLSSKKLLKQYGFPRPPRGSQ 180
 QY 181 AAGGPAKQLPLERVYQETAILKLDHVNKKLIEVLDDPAEDNLYLVFDLLRKGPVME 240
 DB 181 AAGGPAKQLPLERVYQETAILKLDHVNKKLIEVLDDPAEDNLYLVFDLLRKGPVME 240
 QY 241 VPCDKPFSBEQARLYLRDVLGLYLCOKIVHRDIKPSNLLLGDDGHVKIADFGVSNQF 300
 DB 241 VPCDKPFSBEQARLYLRDVLGLYLCOKIVHRDIKPSNLLLGDDGHVKIADFGVSNQF 300
 QY 301 EGNDAQLSSTAGTAPAFMAPEAISDSQSPSGKALDVWATGVTLVCFVYKCPFIDDFILA 360
 DB 301 EGNDAQLSSTAGTAPAFMAPEAISDSQSPSGKALDVWATGVTLVCFVYKCPFIDDFILA 360
 QY 361 LHRKIKNEPVVFPPEPEISEELKDLILKMLDKNPETRIGVDPDIKLPHPWTNKGEPLPSE 420
 DB 361 LHRKIKNEPVVFPPEPEISEELKDLILKMLDKNPETRIGVDPDIKLPHPWTNKGEPLPSE 420
 QY 421 EEHCSSVVEVTEGEVKNVRLIPSWTTVILVKSMLRKSGFNPPPEQARRERSMSAPGNL 480
 DB 421 EEHCSSVVEVTEGEVKNVRLIPSWTTVILVKSMLRKSGFNPPPEQARRERSMSAPGNL 480
 QY 481 LVKEGFGEGGKSPPELPGVQDEAAS 505
 DB 481 LVKEGFGEGGKSPPELPGVQDEAAS 505
 RESULT 2
 ID AAE19161 standard; protein; 505 AA.
 XX
 AC AAE19161;
 XX
 DT 21-MAY-2002 (first entry)
 XX
 DE Human kinase polypeptide (PKIN-19).
 XX
 KW Human; kinase polypeptide; PKIN-19; gene therapy; Addison's disease;
 KW leukaemia; immune disorder; lymphoma; melanoma; developmental disorder;
 KW acquired immunodeficiency syndrome; AIDS; allergy; anaemia; hypertension;
 KW asthma; Crohn's disease; rheumatoid arthritis; bursitis; atherosclerosis;
 KW cirrhosis; hepatitis; psoriasis; Cushing's syndrome; cytostatic; cancer;
 KW cholestasis; cardiact; cardiovascular disorder; Niemann-Pick's disease;
 KW lipid disorder; fatty liver; Gaucher's disease; myocardial infarction;
 KW drug screening; transgenic animal; antiinflammatory; hepatotropic;
 KW hypotensive; anti-HIV; enzyme.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Domain 128..409
 FT Binding-site /note= "Eukaryotic proetin kinase domain"
 FT 485..492
 FT /note= "ATP/GTP-binding site motif A"
 XX
 PN WO200208399-A2.
 XX
 PD 31-JAN-2002.
 XX
 XX 20-JUL-2001; 2001WO-US023092.
 XX
 XX 21-JUL-2000; 2000US-0220038P.
 PR 28-JUL-2000; 2000US-0222112P.
 PR 04-AUG-2000; 2000US-0222831P.
 PR 11-AUG-2000; 2000US-0224729P.
 XX

PA (INCY-) INCYTE GENOMICS INC.
PA (THOR/) THORNTON M.
XX Thornton M, Yue H, Khan FA, Gururajan R, Hafalia AJA, Walia NK;
PI Patterson C, Ramkumar J, Gandhi AR, Policky JL, Baughn MR;
PI Tribouley CM, Bandman O, Nguyen DB, Lu Y, Burford N, Lal P, Ding L;
PI Yao MG, Elliott VS, Recipion SA, Kearney L, Lu DAM, Greenwald SR;
PI Tang YT, Xu Y, Walsh RT, Gietzen KJ, Yang J, Hillman JL;
XX WPI: 2002-206083/26.
DR N-PSDB; AAD30566.
XX
PT New human kinase polypeptide, useful in diagnosis, prevention and
PT treatment of cancer, immune disorder, growth and developmental disorder,
PT cardiovascular disorder and lipid disorder.
XX
PS Claim 1; Page 168-170; 196pp; English.
XX
CC The present invention relates to an isolated human kinase polypeptide
CC (PKIN) or a biologically active/immunogenic fragment of PKIN. PKIN is
CC useful for diagnosing, treating and preventing cancer (e.g., leukaemia,
CC lymphoma, melanoma), an immune disorder (e.g., acquired immunodeficiency
CC syndrome (AIDS), Addison's disease, allergy, anaemia, asthma, Crohn's
CC disease, rheumatoid arthritis), a growth and developmental disorder (e.g.
CC bursitis, cirrhosis, hepatitis, psoriasis, Cushing's syndrome), a
CC cardiovascular disorder (e.g., atherosclerosis, hypertension, myocardial
CC infarction), and a lipid disorder (e.g., fatty liver, cholestasis,
CC Gaucher's disease, Niemann-Pick's disease). PKIN is useful in a number of
CC drug screening techniques and to analyse the proteome of a tissue or cell
CC type. PKIN is useful for creating knockin humanised animals or transgenic
CC animals to model human diseases, in somatic or germline gene therapy, to
CC generate a transcript image of a tissue or cell type, for detecting
CC differences in the chromosomal location due to translocation, inversion,
CC etc., among normal, carrier or affected individuals, and as hybridisation
CC probes for mapping naturally occurring genomic sequences. PKIN is useful
CC in southern or northern analysis, dot blot or other membrane-based
CC technologies, in PCR technologies, in dipstick, pin, multiformat enzyme
CC linked immunosorbent (ELISA)-like assays and in microarrays utilising
CC fluids or tissues from patients to detect altered PKIN expression. The
CC present sequence is human PKIN-19
XX
SQ Sequence 505 AA;
Query Match 99.7%; Score 2626; DB 5; Length 505;
Best Local Similarity 99.8%; Pred. No. 1.6e-235;
Matches 504; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 MEGGPVAVCCQPPRAELVERVAADIVTHLEADGGPEPTRNGVDPPPRARAASVTPGSTSR 60
DB 1 MEGGPVAVCCQPPRAELVERVAADIVTHLEADGGPEPTRNGVDPPPRARAASVTPGSTSR 60
QY 61 LLPARPSLSARKLSQRPAGSYLEAQAGPYATGPASHISPRAWRRPTIESHHVAISDAE 120
DB 61 LLPARPSLSARKLSQRPAGSYLEAQAGPYATGPASHISPRAWRRPTIESHHVAISDAE 120
QY 121 DCVOLNQYKLSQSEIGKAGYVVRVLAAYNESDRHYAMKVLKSKLLKQYGFPRPPRGSQ 180
DB 121 DCVOLNQYKLSQSEIGKAGYVVRVLAAYNESDRHYAMKVLKSKLLKQYGFPRPPRGSQ 180
QY 181 AAQGGPAKQLPLERVQETIAILKLDHVNIVKLIIEVLDDPAEDNLVLFLLRKGPMYE 240
DB 181 AAQGGPAKQLPLERVQETIAILKLDHVNIVKLIIEVLDDPAEDNLVLFLLRKGPMYE 240
QY 241 VPCDKPSEEQARLYRDLVILGLEYLHCQKIVHRDIKPSNLLDGDGHVKIADFGVSNQF 300
DB 241 VPCDKPSEEQARLYRDLVILGLEYLHCQKIVHRDIKPSNLLDGDGHVKIADFGVSNQF 300
QY 301 EGNDAQLSSTAGTAPFAPEAIDSQGSFGSKALDVMATGVTLYCFVYGKCPFFIDDFILA 360
DB 301 EGNDAQLSSTAGTAPFAPEAIDSQGSFGSKALDVMATGVTLYCFVYGKCPFFIDDFILA 360
QY 361 LHRKIKNEPVVFPPEPEISEBELKDLILKMLDKNPETRIGVDDIKLHPWVTYKNGEPELPSE 420
DB 361 LHRKIKNEPVVFPPEPEISEBELKDLILKMLDKNPETRIGVDDIKLHPWVTYKNGEPELPSE 420
QY 421 EEHCSSVVEVEGEVNSVRLIPSWTTTIVLVKSMRLKESFGNPFEPQARRERSMSAFGNL 480
DB 421 EEHCSSVVEVEGEVNSVRLIPSWTTTIVLVKSMRLKESFGNPFEPQARRERSMSAFGNL 480
QY 481 LVXEGFGGKGKSPELPGVQDEAAS 505
DB 481 LVXEGFGGKGKSPELPGVQDEAAS 505
RESULT 3
ADE61939
ID ADE61939 standard; protein; 505 AA.
XX
AC ADE61939;
XX
DT 29-JAN-2004 (first entry)
XX
DB Human Protein NP_115670, SEQ ID NO 7868.
XX
KW Human; pain; neuronal tissue; gene therapy;
KW spinal segmental nerve injury; chronic constriction injury; CCI;
KW spared nerve injury; SNI; Chung.
XX
OS Homo sapiens.
XX
PN WO2003016475-A2.
XX
ED 27-FEB-2003.
XX
XX 14-AUG-2002; 2002WO-US025765.
XX
PF 14-AUG-2001; 2001US-0312147P.
XX
PR 01-NOV-2001; 2001US-0346382P.
XX
PR 26-NOV-2001; 2001US-0333347P.
XX
PA (GSHO) GEN HOSPITAL CORP.
PA (FARB) BAYER AG.
XX
PI Woolf C, D'urso D, Befort K, Costigan M;
XX
DR WPI; 2003-269312/26.
XX
DR GENBANK; NP_115670.
XX
PT New composition comprising two or more isolated polypeptides, useful for
PT preparing a medicament for treating pain in an animal.
XX
XX Claim 1; Page; 1017pp; English.
CC
CC The invention discloses a composition comprising two or more isolated rat
CC or human polynucleotides or a polynucleotide which represents a fragment,
CC derivative or allelic variation of the nucleic acid sequence. Also
CC claimed are a vector comprising the novel polynucleotide, a host cell
CC comprising the vector, a method for identifying a nucleotide sequence
CC which is differentially regulated in an animal subjected to pain and a
CC kit to perform the method, an array, a method for identifying an agent
CC that increases or decreases the expression of the polynucleotide sequence
CC that is differentially expressed in neuronal tissue of a first animal
CC subjected to pain, a method for identifying a compound which regulates
CC the expression of a polynucleotide sequence which is differentially
CC expressed in an animal subjected to pain, a method for identifying a
CC compound that regulates the activity of one or more of the
CC polynucleotides, a method for producing a pharmaceutical composition, a
CC method for identifying a compound or small molecule that regulates the
CC activity in an animal of one or more of the polypeptides given in the
CC specification, a method for identifying a compound useful in treating
CC pain and a pharmaceutical composition comprising the one or more
CC polypeptides or their antibodies. The polynucleotide or the compound that
CC modulates its activity is useful for preparing a medicament for treating
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC therapy). The sequence presented is a human protein (shown in Table 2 of

CC the specification) which is differentially expressed during pain. Note:
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 505 AA;
Query Match 99.7%; Score 2626; DB 7; Length 505;
Best Local Similarity 99.8%; Pred. No. 1.6e-235;
Matches 504; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 MEGGPVCCQDPRAELVERVAALDVTHLEADGGPEPTNGVDPPPRARAASVIPGSTSR 60
DB 1 MEGGPVCCQDPRAELVERVAALDVTHLEADGGPEPTNGVDPPPRARAASVIPGSTSR 60
QY 61 LLPAPSLARKLSLOERPAAGSYLEAQAGYATGPASHISPRAWRPPTTIESHVAISDAE 120
DB 61 LLPAPSLARKLSLOERPAAGSYLEAQAGYATGPASHISPRAWRPPTTIESHVAISDAE 120
QY 121 DCVQLNQYKLOSEIGKYGAVVRLAYNESEDHRYAMKVLKSKLLKQYGFPRPPRGSQ 180
DB 121 DCVQLNQYKLOSEIGKYGAVVRLAYNESEDHRYAMKVLKSKLLKQYGFPRPPRGSQ 180
QY 181 AAQGGPAKQLPLERVYQETAILKKLDHVNKKLIEVLDDPAEDNLYLVFDLLRKGPMVE 240
DB 181 AAQGGPAKQLPLERVYQETAILKKLDHVNKKLIEVLDDPAEDNLYLVFDLLRKGPMVE 240
QY 241 VPCDKPFSEEQARLYLRDVLIGLYLHCQKIVHRDIKPSNLLGDDGHVKIADFGVSNQF 300
DB 241 VPCDKPFSEEQARLYLRDVLIGLYLHCQKIVHRDIKPSNLLGDDGHVKIADFGVSNQF 300
QY 301 EGNDAQLSSTAGTAPFAPEAISDSGQSGKALDVMATGVTLYCFYVKGKCPIDDFILA 360
DB 301 EGNDAQLSSTAGTAPFAPEAISDSGQSGKALDVMATGVTLYCFYVKGKCPIDDFILA 360
QY 361 LHRKIKNEPVVFPPEEISEELKDLILKMLDKNPETRIGVDPDIKLPWTKNGEELPSE 420
DB 361 LHRKIKNEPVVFPPEEISEELKDLILKMLDKNPETRIGVDPDIKLPWTKNGEELPSE 420
QY 421 EECSCVVEVTEGEVKNRSLIPSWTIVLVKSLMRKRSFGNPPPOARREERSMSAPGNL 480
DB 421 EECSCVVEVTEGEVKNRSLIPSWTIVLVKSLMRKRSFGNPPPOARREERSMSAPGNL 480
QY 481 LVKEGFGGKSPELQVQDEAAS 505
DB 481 LVKEGFGGKSPELQVQDEAAS 505

RESULT 4
ADE61943
XX ADE61943 standard; protein; 505 AA.
XX AC ADE61943;
XX XX
XX DT 29-JAN-2004 (first entry)
XX DE Human Protein NP_115670, SEQ ID NO 7872.
XX KW Human; pain; neuronal tissue; gene therapy;
XX KW spinal segmental nerve injury; chronic constriction injury; CCI;
XX KW spared nerve injury; SNI; Chung.
XX OS Homo sapiens.
XX PN WC2003016475-A2.
XX PD 27-FEB-2003.
XX PF 14-AUG-2002; 2002WO-US025765.
XX PR 14-AUG-2001; 2001US-0312147P.
XX PR 01-NOV-2001; 2001US-0346362P.
XX PR 26-NOV-2001; 2001US-0333347P.

XX (GEO) GEN HOSPITAL CORP.
PA (FARB) BAYER AG.
XX
XX PI Woolf C, D'urso D, Befort K, Costigan M;
XX
XX WPI; 2003-268312/26.
DR GENBANK, NP_115670.
XX
XX New composition comprising two or more isolated polypeptides, useful for
PT preparing a medicament for treating pain in an animal.
PS
XX Claim 1; Page; 1017pp; English.
CC The invention discloses a composition comprising two or more isolated rat
CC or human polynucleotides or a polynucleotide which represents a fragment,
CC derivative or allelic variation of the nucleic acid sequence. Also
CC claimed are a vector comprising the novel polynucleotide, a host cell
CC comprising the vector, a method for identifying a nucleotide sequence
CC which is differentially regulated in an animal subjected to pain and a
CC kit to perform the method, an array, a method for identifying an agent
CC that increases or decreases the expression of the polynucleotide sequence
CC that is differentially expressed in neuronal tissue of a first animal
CC subjected to pain, a method for identifying a compound which regulates
CC the expression of a polynucleotide sequence which is differentially
CC expressed in an animal subjected to pain, a method for identifying a
CC compound that regulates the activity of one or more of the
CC polynucleotides, a method for producing a pharmaceutical composition, a
CC method for identifying a compound or small molecule that regulates the
CC activity in an animal of one or more of the polypeptides given in the
CC specification, a method for identifying a compound useful in treating
CC pain and a pharmaceutical composition comprising the one or more
CC polypeptides or their antibodies. The polynucleotide or the compound that
CC modulates its activity is useful for preparing a medicament for treating
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC therapy). The sequence presented is a human protein (shown in Table 2 of
CC the specification) which is differentially expressed during pain. Note:
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX Sequence 505 AA;

Query Match 99.7%; Score 2626; DB 7; Length 505;
Best Local Similarity 99.8%; Pred. No. 1.6e-235;
Matches 504; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 MEGGPVCCQDPRAELVERVAALDVTHLEADGGPEPTNGVDPPPRARAASVIPGSTSR 60
DB 1 MEGGPVCCQDPRAELVERVAALDVTHLEADGGPEPTNGVDPPPRARAASVIPGSTSR 60
QY 61 LLPAPSLARKLSLOERPAAGSYLEAQAGYATGPASHISPRAWRPPTTIESHVAISDAE 120
DB 61 LLPAPSLARKLSLOERPAAGSYLEAQAGYATGPASHISPRAWRPPTTIESHVAISDAE 120
QY 121 DCVQLNQYKLOSEIGKYGAVVRLAYNESEDHRYAMKVLKSKLLKQYGFPRPPRGSQ 180
DB 121 DCVQLNQYKLOSEIGKYGAVVRLAYNESEDHRYAMKVLKSKLLKQYGFPRPPRGSQ 180
QY 181 AAQGGPAKQLPLERVYQETAILKKLDHVNKKLIEVLDDPAEDNLYLVFDLLRKGPMVE 240
DB 181 AAQGGPAKQLPLERVYQETAILKKLDHVNKKLIEVLDDPAEDNLYLVFDLLRKGPMVE 240
QY 241 VPCDKPFSEEQARLYLRDVLIGLYLHCQKIVHRDIKPSNLLGDDGHVKIADFGVSNQF 300
DB 241 VPCDKPFSEEQARLYLRDVLIGLYLHCQKIVHRDIKPSNLLGDDGHVKIADFGVSNQF 300
QY 301 EGNDAQLSSTAGTAPFAPEAISDSGQSGKALDVMATGVTLYCFYVKGKCPIDDFILA 360
DB 301 EGNDAQLSSTAGTAPFAPEAISDSGQSGKALDVMATGVTLYCFYVKGKCPIDDFILA 360
QY 361 LHRKIKNEPVVFPPEEISEELKDLILKMLDKNPETRIGVDPDIKLPWTKNGEELPSE 420

PR 13-MAR-2001; 2001US-0275579P.
PR 13-MAR-2001; 2001US-0275601P.
PR 14-MAR-2001; 2001US-0276000P.
PR 16-MAR-2001; 2001US-0276776P.
PR 19-MAR-2001; 2001US-0276994P.
PR 20-MAR-2001; 2001US-0277339P.
PR 20-MAR-2001; 2001US-0277321P.
PR 20-MAR-2001; 2001US-0277327P.
PR 21-MAR-2001; 2001US-0277791P.
PR 22-MAR-2001; 2001US-0277738P.
PR 23-MAR-2001; 2001US-0278152P.
PR 26-MAR-2001; 2001US-0278894P.
PR 27-MAR-2001; 2001US-0278999P.
PR 27-MAR-2001; 2001US-0279036P.
PR 28-MAR-2001; 2001US-0279344P.
PR 30-MAR-2001; 2001US-0277338P.
PR 30-MAR-2001; 2001US-0279995P.
PR 30-MAR-2001; 2001US-0280233P.
PR 02-APR-2001; 2001US-0280802P.
PR 02-APR-2001; 2001US-0280822P.
PR 02-APR-2001; 2001US-0280900P.
PR 04-APR-2001; 2001US-0281194P.
PR 13-APR-2001; 2001US-0283675P.
PR 30-APR-2001; 2001US-0287424P.
PR 02-MAY-2001; 2001US-0288066P.
PR 03-MAY-2001; 2001US-0288342P.
PR 03-MAY-2001; 2001US-0288528P.
PR 15-MAY-2001; 2001US-0291190P.
PR 16-MAY-2001; 2001US-0291099P.
PR 16-MAY-2001; 2001US-0291240P.
PR 30-MAY-2001; 2001US-0294485P.
PR 31-MAY-2001; 2001US-0294889P.
PR 31-MAY-2001; 2001US-0294899P.
PR 18-JUN-2001; 2001US-0299027P.
PR 19-JUN-2001; 2001US-0299303P.
PR 19-JUN-2001; 2001US-0299310P.
PR 10-JUL-2001; 2001US-0304354P.
PR 31-JUL-2001; 2001US-0309198P.
PR 16-AUG-2001; 2001US-0312903P.
PR 10-SEP-2001; 2001US-0318462P.
PR 12-SEP-2001; 2001US-0318770P.
PR 27-SEP-2001; 2001US-0325430P.
PR 27-SEP-2001; 2001US-0325681P.
PR 18-OCT-2001; 2001US-0330380P.
PR 31-OCT-2001; 2001US-0335301P.
PR 14-NOV-2001; 2001US-0332172P.
PR 14-NOV-2001; 2001US-0332271P.
PR 14-NOV-2001; 2001US-0332272P.
PR 14-NOV-2001; 2001US-0333184P.
PR 14-NOV-2001; 2001US-0333272P.
PR 21-NOV-2001; 2001US-0332094P.
PR 03-DEC-2001; 2001US-0337426P.
PR 03-DEC-2001; 2001US-0338092P.
PR 04-DEC-2001; 2001US-0337185P.
PR 03-JAN-2002; 2002US-0345705P.
PR 07-MAR-2002; 2002US-00092900.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Padigar M, Spytek KA, Shenoy SG, Taupier RJ, Pena CEA, Li L;
PI Zehrusen BD, Gusev V, Ji W, Gorman I, Miller CE, Kekuda R;
PI Pattarajan M, Gangolli E, Vernet CAM, Guo X, Tchernev V;
PI Fernandes ER, Casman SJ, Malyankar UM, Gerlach V, Liu Y, Anderson D;
PI Spaderna SK, Catterton E, Burgess C, Leite M, Zhong H, Alsobrook JP;
PI Lepley DW, Rieger DK;
XX
XX WPI; 2002-723332/78.
DR N-PSDB; ABX97035.
DR
XX
XX NOVX polypeptides and polynucleotides, useful for preventing or treating
PT a disorder associated with aberrant NOVX expression or activity e.g.,
PT cancer, hypertension, atherosclerosis, cardiomyopathy or bronchial
PT asthma.

XX
PS Claim 1; Page 139; 1103pp; English.
XX
CC This invention describes novel human NOVX polypeptides which have
CC cytostatic, cardiant, antiarteriosclerotic, antiasthmatic and hypotensive
CC activity. Pharmaceutical compositions comprising the NOVX proteins or
CC nucleic acid molecules or NOVX antibodies are useful for preventing or
CC treating a disorder associated with aberrant NOVX expression or activity
CC e.g. cancer, hypertension, atherosclerosis, cardiomyopathy or bronchial
CC asthma. The products of the invention can be used for gene therapy or in
CC a vaccine. ABU65041-ABU65218 represent the NOVX polypeptides encoded by
CC ABX97008-ABX97185
XX
SQ Sequence 505 AA;

Query Match 99.1%; Score 2611; DB 5; Length 505;
Best Local Similarity 99.4%; Pred. No. 3.9e-234;
Matches 502; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 MEGGPAVCCQDPRAELVERVAALDVTHLEADGGPEPTRNGVDPPPARAASVIPGSTSR 60
DB 1 MEGGPAVCCQDPRAELVERVAALDVTHLEADGGPEPTRNGVDPPPARAASVIPGSTSR 60
QY 61 LLPAPESLSARKLSLQERPAGSYLEAQAGFYATGPASHISPRAWRPTTIESHVAISDAE 120
DB 61 LLPAPESLSARKLSLQERPAGSYLEAQAGFYATGPASHISPRAWRPTTIESHVAISDAE 120
QY 121 DCVQLNQYLQSEIGHGKAGVVRVLAINESEDRHYAMKVLKKKLLKQYGFPRPPRGSQ 180
DB 121 DCVQLNQYLQSEIGHGKAGVVRVLAINESEDRHYAMKVLKKKLLKQYGFPRPPRGSQ 180
QY 181 AAOGGPAKOLLPLERYVQEIALLKLDHNVVVKLIEVLDDPAEDNLYLVFDLLRKGPVME 240
DB 181 AAOGGPAKOLLPLERYVQEIALLKLDHNVVVKLIEVLDDPAEDNLYLVFDLLRKGPVME 240
QY 241 VPCDKPFSEEQARLYLRDVLGLYLELHCQKIVHRDIKPSNLLLGDDGHVKIADFGVSNQF 300
DB 241 VPCDKSFSEEQARLYLRDVLGLYLELHCQKIVHRDIKPSNLLLGDDGHVKIADFGVSNQF 300
QY 301 EGNDAQLSSTAGTAPMAPEAISDSGSGKALDYWATGVTLYCFYVGCPIDDFILA 360
DB 301 EGNDAQLSSTAGTAPMAPEAISDSGSGKALDYWATGVTLYCFYVGCPIDDFILA 360
QY 361 LHRKIKNEPVVFPPEPEISEELKDLILKMLDKNKPETRIQVDPDIKLPWTKNGEELPSE 420
DB 361 LHRKIKNEPVVFPPEPEISEELKDLILKMLDKNKPETRIQVDPDIKLPWTKNGEELPSE 420
QY 421 EEHCVVVEVTEGEVKNSVRLIPSWTTVLVKSMLRKRSFGNPFEPQARRERSMSAPGNL 480
DB 421 EEHCVVVEVTEGEVKNSVRLIPSWTTVLVKSMLRKRSFGNPFEPQARRERSMSAPGNL 480
QY 481 LVKEGFGEGGKSPQLPGVQDEAAS 505
DB 481 LVKEGFGEGGKSPQLPGVQDEAAS 505

RESULT 7
AAU03510
ID AAU03510 standard; protein; 513 AA.
XX
AC AAU03510;
XX
DT 12-SEP-2001 (first entry)
XX
XX Human protein kinase #10.
XX
XX Human; protein kinase; PTK; STK; cancer; cardiovascular disease;
KW metabolic disorder; immune related disease; neurological disorder;
KW neurodegenerative disorder; inflammatory disorder; infectious disease;
KW reproductive disorder.
XX
OS Homo sapiens.
XX

FN W0200138503-A2.
XX
PD 31-MAY-2001.
XX
PF 22-NOV-2000; 2000WO-US032085.
XX
PR 24-NOV-1999; 99US-0167482P.
XX
PA (SUCGE-) SUGEN INC.
XX
PI Plowman GD, Whyte D, Manning G, Sudarsanam S, Martinez R;
PI Flanagan P, Clary D;
XX WPI: 2001-343950/36.
DR N-PSDB; AAS06710.
XX
FT Nucleic acids encoding human kinase polypeptides, useful for preventing
FT diagnosing and/or treating e.g. cancer, immune, cardiovascular and
FT neuronal-associated diseases, and microbial infections.
XX
PS Claim 7; Fig 2; 433pp; English.
XX
CC AAU03501-AAU03557 represent novel human protein kinases #1-57. The novel
CC protein kinases have been identified as members of the tyrosine or
CC serine/threonine kinase (PTK and STK) families. The polynucleotides
CC encoding protein kinases and the polypeptides may be used in the
CC prevention, diagnosis and treatment of diseases associated with
CC inappropriate kinase expression. For example, they may be used to treat
CC cancers (especially cancers of haematopoietic origin), cardiovascular
CC disease (e.g. atherosclerosis), metabolic disorders (e.g. diabetes),
CC immune related diseases (e.g. rheumatoid arthritis), neurological
CC disorders (e.g. schizophrenia), neurodegenerative disorders (e.g.
CC Parkinson's disease), inflammatory disorders (e.g. asthma), infectious
CC disease (e.g. HIV) and reproductive disorders (e.g. infertility).
CC Additionally, polynucleotides encoding protein kinases may be used for
CC gene therapy and as DNA probes in diagnostic assays. The protein kinase
CC polypeptides may be used as antigens in the production of antibodies
CC against the protein kinases and in assays to identify modulators of
CC protein kinase expression and activity
XX
SQ Sequence 513 AA;
Query Match 98.5%; Score 2595; DR 4; Length 513;
Best Local Similarity 97.9%; Pred. No. 1.3e-232;
Matches 503; Conservative 0; Mismatches 1; Indels 10; Gaps 2;
Qy 1 MEGGPAVCCQDPRAELVERVAIDVTHLEADGGPEPTRNGVDPPRARAASVIPGSTSR 60
Db 1 MEGGPAVCCQDPRAELVERVAIDVTHLEADGGPEPTRNGVDPPRARAASVIPGSTSR 60
Qy 61 LLPARPSLSARKLSLOERPAGSYLEAQAGPYATGPASHISPRWRPTIESHHVAISDAE 120
Db 61 LLPARPSLSARKLSLOERPAGSYLEAQAGPYATGPASHISPRWRPTIESHHVAISDAE 120
Qy 121 DCVQLNOYKLOSEIGK-----GAYGVVRLAYNESEDRHYAMKVLKSKKLKQYGF 171
Db 121 DCVQLNQYKLOSEIGKVGCLTDAYLQAGVGVRLAYNESEDRHYAMKVLKSKKLKQYGF 180
Qy 172 RRPFRGSQAAGQGPAPKQLPLERYQIEAILKKLDHNVVVKLIEVLDDPAEDNLYLVFD 231
Db 181 RRPFRGSQAAGQGPAPKQLPLERYQIEAILKKLDHNVVVKLIEVLDDPAEDNLYLV-D 239
Qy 232 LLRKGPMVEPCDFPSEQARLYLRDVLGLEYLHCOKIVHRDIKPSNLLLGDDGHVKI 291
Db 240 LLRKGPMVEPCDFPSEQARLYLRDVLGLEYLHCOKIVHRDIKPSNLLLGDDGHVKI 299
Qy 292 ADFGVSNQFEGNDQALSTAGTAPFAPPEATSDSQSGKALDVMATGVTLYCFVYKGC 351
Db 300 ADFGVSNQFEGNDQALSTAGTAPFAPPEATSDSQSGKALDVMATGVTLYCFVYKGC 359
Qy 352 PFIDDFILALHRKIKNEPVVFPPEPSEELKDLILKMLDKNPETRIGVPDIKLPWVTK 411
Db 360 PFIDDFILALHRKIKNEPVVFPPEPSEELKDLILKMLDKNPETRIGVPDIKLPWVTK 419

Qy 412 NGEELPSEEEHCSSVVEVTEGEVKNSVRLIPSWTTVLVKMLRKESFGNPFEPQARREE 471
Db 420 NGEELPSEEEHCSSVVEVTEGEVKNSVRLIPSWTTVLVKMLRKESFGNPFEPQARREE 479
Qy 472 RMSAPGNLLVKEGFGEGGKSPGLPGVQDEAAS 505
Db 480 RMSAPGNLLVKEGFGEGGKSPGLPGVQDEAAS 513
RESULT 8
AAE04361
ID AAE04361 standard; protein; 513 AA.
XX AC AAE04361;
XX DT 04-SEP-2001 (first entry)
XX DE Human kinase (PKIN)-2.
XX KW Human kinase; PKIN-2; therapy; immune disorder; Addison's disease; AIDS;
KW acquired immune deficiency syndrome; growth and developmental disorder;
KW arteriosclerosis; mixed connective tissue disease; MCTD; adenocarcinoma;
KW leukaemia; cardiovascular disease; myocardial infarction; hypertension;
KW lipid disorder; cancer; fatty liver; cholestasis; transgenic animal;
KW gene therapy; antiallergic; antiasthmatic; antithyroid; dermatological;
KW antidiabetic; nephrotrophic; anticancer; antiarthritic; antirheumatic;
KW antipsoriatic; neuroprotective; cytostatic; hepatotropic; osteopathic;
KW vasotropic; antianginal; anorectic.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
XX Region 1..127
FT /note= "Kinase protein beta"
FT Domain 130..408
FT /note= "protein kinase domain"
FT Domain 145..417
FT /note= "Eukaryotic protein kinase domain"
FT Domain 273..291
FT /note= "Tyrosine kinase catalytic domain"
FT Domain 320..330
FT /note= "Tyrosine kinase catalytic domain"
FT Domain 342..364
FT /note= "Tyrosine kinase catalytic domain"
FT Binding-site 493..500
FT /note= "ATP/GTP-binding site motif A (P-loop)"
XX W0200146397-A2.
XX
PD 28-JUN-2001.
XX
PF 20-DEC-2000; 2000WO-US035304.
XX
PR 23-DEC-1999; 99US-0172066P.
PR 14-JAN-2000; 2000US-0176107P.
PR 21-JAN-2000; 2000US-017731P.
PR 28-JAN-2000; 2000US-0178573P.
XX
PA (INCY-) INCYTE GENOMICS INC.
XX
PI Yang J, Baughn MR, Burford N, Au-Young J, Lu DAM, Reddy R;
PI Yue H, Yao MG, Lal P, Khan FA;
XX
DR WPI: 2001-418059/44.
DR N-PSDB; AAD08635.
XX
FT Novel human kinase proteins (PKIN) useful for diagnosing, treating,
FT preventing immune disorders, cardiovascular diseases and disorders
FT affecting growth and development associated with abnormal expression of
FT PKIN.
XX
PS Claim 1; Page 104-105; 128pp; English.

D6	1	MEGPAVCCQDPRAELVERAAIDVTHLEBADGGPEPTRNGVDPPPPARAASVPIPGTSR	60
QY	61	LLPARPSLSARKLSLQERPAGSYLEAQAGPYATGPASHISPRAMRPTIESHHVAISDAE	120
D6	61	LLPARPSLSARKLSLQERPAGSYLEAQAGPYATGPASHISPRAMRPTIESHHVAISDAE	120
QY	121	DCVQLNQYKLOSEIGKYGAVVRLAYNESEDRHYAMKVLKKLLKQYGFRRPPRGSG	180
D6	121	DCVQLNQYKLOSEIGKYGAVVRLAYNESEDRHYAMKVLKKLLKQYGFRRPPRGSG	180
QY	181	AAQGGPAKQLPLERVYQEIALLKKLDHVVNVVKLIEVLDDPAEDNLYLVFLLLRKPWME	240
D6	181	AAQGGPAKQLPLERVYQEIALLKKLDHVVNVVKLIEVLDDPAEDNLYLVFLLLRKPWME	240
QY	241	VPCDKPFSEEQARLYLRDVLGLEYLHCQKIVHRDIKPSNLLGDDGHVKIADFGVSNQF	300
D6	241	VPCDKPFSEEQARLYLRDVLGLEYLHCQKIVHRDIKPSNLLGDDGHVKIADFGVSNQF	300
QY	301	EGNDAQLSSTAGTAFAPMAPEAISDSGQSFSGKALDVWATGVTLYCFVYGKCPFFIDDFILA	360
D6	301	EGNDAQLSSTAGTAFAPMAPEAISDSGQSFSGKALDVWATGVTLYCFVYGKCPFFIDDFILA	360
QY	361	LHRKIKNEPVVPPEPEISEELKDILKMLDKNPETRIGVDPDIKLHPWVTNKGEEPLPSE	420
D6	361	LHRKIKNEPVVPPEPEISEELKDILKMLDKNPETRIGVDPDIKLHPWVTNKGEEPLPSE	420
QY	421	EEHCSVVEVTEGEVKNVRLIPSWTTVILVKSMLRKRSFGNPFEPQARRERSMSAPGNL	480
D6	421	EEHCSVVEVTEGEVKNVRLIPSWTTVILVKSMLRKRSFGNPFEPQARRERSMSAPGNL	480
QY	481	LVKEGFGGKSPELPGVQ	499
D6	481	LVKEGFGGKSPELPGVQ	499
RESULT 10			
ABU65069	ID	ABU65069 standard; protein; 543 AA.	
XX	AC	ABU65069;	
XX	XX	20-MAY-2003 (first entry)	
DT	XX	Human NOV15c protein.	
DE	XX	NOVX; cytostatic; cardiant; antiarteriosclerotic; antiasthmatic; cancer;	
KW	KW	hypotensive; cardiomyopathy; bronchial asthma; gene therapy; vaccine;	
KW	KW	human.	
XX	OS	Homo sapiens.	
XX	XX	WO200272757-A2.	
XX	XX	19-SEP-2002.	
XX	XX	08-MAR-2002; 2002WO-US006908.	
PR	PR	08-MAR-2001; 2001US-0274101P.	
PR	PR	08-MAR-2001; 2001US-0274194P.	
PR	PR	08-MAR-2001; 2001US-0274281P.	
PR	PR	08-MAR-2001; 2001US-0274322P.	
PR	PR	09-MAR-2001; 2001US-0274849P.	
PR	PR	12-MAR-2001; 2001US-0275235P.	
PR	PR	13-MAR-2001; 2001US-0275578P.	
PR	PR	13-MAR-2001; 2001US-027579P.	
PR	PR	13-MAR-2001; 2001US-0275601P.	
PR	PR	14-MAR-2001; 2001US-0276000P.	
PR	PR	16-MAR-2001; 2001US-0276776P.	
PR	PR	19-MAR-2001; 2001US-0276994P.	
PR	PR	20-MAR-2001; 2001US-0277239P.	
PR	PR	20-MAR-2001; 2001US-0277321P.	
PR	PR	20-MAR-2001; 2001US-0277327P.	
PR	PR	21-MAR-2001; 2001US-0277791P.	

PR	22-MAR-2001; 2001US-0277833P.	
PR	23-MAR-2001; 2001US-0278152P.	
PR	26-MAR-2001; 2001US-0278894P.	
PR	27-MAR-2001; 2001US-0278999P.	
PR	27-MAR-2001; 2001US-0279036P.	
PR	28-MAR-2001; 2001US-0279344P.	
PR	30-MAR-2001; 2001US-0277338P.	
PR	30-MAR-2001; 2001US-0279995P.	
PR	30-MAR-2001; 2001US-0280233P.	
PR	02-APR-2001; 2001US-0280802P.	
PR	02-APR-2001; 2001US-0280822P.	
PR	02-APR-2001; 2001US-0280900P.	
PR	04-APR-2001; 2001US-0281194P.	
PR	13-APR-2001; 2001US-0283675P.	
PR	30-APR-2001; 2001US-0287424P.	
PR	02-MAY-2001; 2001US-0288066P.	
PR	03-MAY-2001; 2001US-0288342P.	
PR	03-MAY-2001; 2001US-0288528P.	
PR	15-MAY-2001; 2001US-0291190P.	
PR	16-MAY-2001; 2001US-0291099P.	
PR	16-MAY-2001; 2001US-0291240P.	
PR	30-MAY-2001; 2001US-029485P.	
PR	31-MAY-2001; 2001US-0294889P.	
PR	31-MAY-2001; 2001US-0294899P.	
PR	18-JUN-2001; 2001US-0299027P.	
PR	19-JUN-2001; 2001US-0299303P.	
PR	19-JUN-2001; 2001US-0299310P.	
PR	10-JUL-2001; 2001US-0304354P.	
PR	31-JUL-2001; 2001US-0309198P.	
PR	16-AUG-2001; 2001US-0312903P.	
PR	10-SEP-2001; 2001US-0318462P.	
PR	12-SEP-2001; 2001US-0318770P.	
PR	27-SEP-2001; 2001US-0325430P.	
PR	27-SEP-2001; 2001US-0325681P.	
PR	18-OCT-2001; 2001US-0330380P.	
PR	31-OCT-2001; 2001US-0335301P.	
PR	14-NOV-2001; 2001US-0332172P.	
PR	14-NOV-2001; 2001US-0332271P.	
PR	14-NOV-2001; 2001US-0332727P.	
PR	14-NOV-2001; 2001US-0333184P.	
PR	14-NOV-2001; 2001US-0333272P.	
PR	21-NOV-2001; 2001US-0332094P.	
PR	03-DEC-2001; 2001US-0337426P.	
PR	03-DEC-2001; 2001US-0338092P.	
PR	04-DEC-2001; 2001US-0337185P.	
PR	03-JAN-2002; 2002US-0345705P.	
PR	07-MAR-2002; 2002US-00092900.	
XX		
XX	(CURA-) CURAGEN CORP.	
PI	Padigaru M, Spytek KA, Shenoy SG, Taupier RJ, Pena CEA, Li L;	
PI	Zerhusen BD, Gusev V, Ji W, Gorman L, Miller CE, Kekuda R;	
PI	Patturajan M, Gangolli E, Vernet CAM, Guo X, Tchernev V;	
PI	Fernandes ER, Casman S, Malyankar UM, Gerlach V, Liu Y, Anderson D;	
PI	Spaderna SK, Catterton E, Burgess C, Leite M, Zhong H, Alsobrook JP;	
PI	Lepley DM, Rieger DK;	
XX		
XX	WPI; 2002-723332/78.	
DR	N-PSDB; ABX97036.	
XX		
PT	NOVX polypeptides and polynucleotides, useful for preventing or treating	
PT	a disorder associated with aberrant NOVX expression or activity e.g.,	
PT	cancer, hypertension, atherosclerosis, cardiomyopathy or bronchial	
PT	asthma.	
XX		
PS	Claim 1; Page 140; 1103pp; English.	
XX		
CC	This invention describes novel human NOVX polypeptides which have	
CC	cytostatic, cardiant, antiarteriosclerotic, antiasthmatic and hypotensive	
CC	activity. Pharmaceutical compositions comprising the NOVX proteins or	
CC	nucleic acid molecules or NOVX antibodies are useful for preventing or	
CC	treating a disorder associated with aberrant NOVX expression or activity	
CC	e.g. cancer, hypertension, atherosclerosis, cardiomyopathy or bronchial	

DR WPI; 2002-723332/78.
DR N-PSDB; ABX97034.
XX
XX NOVX polypeptides and polynucleotides, useful for preventing or treating
PT a disorder associated with aberrant NOVX expression or activity e.g.,
PT cancer, hypertension, atherosclerosis, cardiomyopathy or bronchial
PT asthma.
XX
XX
PS Claim 1; Page 139; 1103pp; English.
XX
XX This invention describes novel human NOVX polypeptides which have
CC cyrostatic, cardiant, antiarteriosclerotic, antiasthmatic and hypotensive
CC activity. Pharmaceutical compositions comprising the NOVX proteins or
CC nucleic acid molecules or NOVX antibodies are useful for preventing or
CC treating a disorder associated with aberrant NOVX expression or activity
CC e.g. cancer, hypertension, atherosclerosis, cardiomyopathy or bronchial
CC asthma. The products of the invention can be used for gene therapy or in
CC a vaccine. ABU65041-ABU65218 represent the NOVX polypeptides encoded by
CC ABX97008-ABX97185
XX
XX Sequence 503 AA;
XX
Query Match 97.2%; Score 2560; DB 5; Length 503;
Best Local Similarity 98.2%; Pred. No. 2.2e-229;
Matches 496; Conservative 2; Mismatches 5; Indels 2; Gaps 2;
Qy 1 MEGPVCVCCQDPRAELVERVAADIVTHLEADGGPEPTRNGVDPPPRARAASVIPGSTSR 60
Db 1 MEGPVCVCCQDPRAELVERVAADIVTHLEADGGPEPTRNGVDPPPRARAASVIPGSTSR 60
Qy 61 LLPARPSLSARKLSLOERPAGSYLEAQAGPYATGPASHISPRAWRPTIESHHVAISDAE 120
Db 61 LLPARPSLSARKLSLOERPAGSYLEAQAGPYATGPASHISPRAWRPTIESHHVAISDAE 120
Qy 121 DCVOLNQYKQSEIGKGYGVVRLAYNESEDRHYAMKVLSSKKLLKXQYGFPRPPRGSQ 180
Db 121 DCVOLNQYKQSEIGKGYGVVRLAYNESEDRHYAMKVLSSKKLLKXQYGFPRPPRGSQ 180
Qy 181 AAQGGPAKQLPLERVYQETAILKKLDHVNKKLIEVLDDPAEDNLKLVFLLRKGPMVE 240
Db 181 AAQGGPAKQLPLERVYQETAILKKLDHVNKKLIEVLDDPAEDNLKLVFLLRKGPMVE 240
Qy 241 VPCDKPSEEARLYLRDVLGLYHCQKIVHRDIKPSNLLLGDDGHVKIADFGVSNQF 300
Db 240 VPCDKPSEEARLYLRDVLGLYHCQKIVHRDIKPSNLLLGDDGHVKIADFGVSNQF 299
Qy 301 EGNDALSSSTAGTAPFAPEAISDSGQSFSGKALDVWATGVTLYCFVYGKCPFFIDDFILA 360
Db 300 EGNDALSSSTAGTAPFAPEAISDSGQSFSGK-LDVWATGVTLYCFVYGKCPFFIDDFILA 358
Qy 361 LHRKIKNEPVVFPPEPEISEELKOLILKMLDKNPETRIGVPDIKLVHPWVTKNGEPLPSE 420
Db 359 LHRKIKNEPVVFPPEPEISEELKOLILKMLDKNPETRIGVPDIKLVHPWVTKNGEPLPSE 418
Qy 421 EEHCSVVEVTGEVKNVRLIPSTVTIVLKSMLKRSGFNGFPFQARREERSAPGNL 480
Db 419 EEHCSVVEVTGEVKNVRLIPSTVTIVLKSMLKRSGFNGFPFQARREERSAPGNL 478
Qy 481 LVKSGFEGGKSPELPGVQDEAAS 505
Db 479 LVKSGFEGGKSPELPGVQDEAAS 503
RESULT 12
ID ABB83055
XX ABB83055 standard; protein; 505 AA.
AC ABB83055;
XX
XX 27-AUG-2002 (first entry)
XX
XX Rat amino acid sequence related to the kinase protein of the invention.
XX

KW Rat; kinase protein; calcium/calmodulin-dependant protein kinase; eye;
KW retinoblastoma; brain; kinase modulator.
XX
XX Rattus norvegicus.
XX
XX WO200224920-A2.
XX
XX 28-MAR-2002.
XX
XX 19-SEP-2001; 2001WO-US029161.
XX
XX 19-SEP-2000; 2000US-0233493P.
XX 13-NOV-2000; 2000US-0247031P.
XX 06-DEC-2000; 2000US-00729995.
XX
XX (PEKE) PE CORP NY.
XX
XX Beasley EM, Wei M, Bonazzi VR, Sanders R, Di Francesco V;
XX WPI; 2002-404955/43.
XX
XX Novel peptide designated as human kinase useful as target for diagnosing
PT a disease or predisposition to the disease mediated by the peptide.
XX
XX Disclosure; Page 88-89; 89pp; English.
XX
XX The invention relates to an isolated peptide designated human kinase
CC (HK), that has homology to members of the calcium/calmodulin-dependent
CC protein kinase kinase subfamily. The mechanism of action of the protein
CC of the invention is that of a kinase modulator. The human kinase of the
CC invention is useful for creating a pharmaceutical composition for
CC treating a disease or condition mediated by the human kinase. HK is also
CC useful to provide a target for diagnosing a disease or predisposition to
CC disease mediated by HK, and is also useful in pharmacogenic analysis. HK
CC is useful for treating a disorder characterised by absence of
CC inappropriate or unwanted expression of HK, also as an immunogen to raise
CC antibodies by administering HK to a mammalian organism e.g. rat, rabbit
CC or mouse. Nucleic acids of the invention are useful as hybridisation
CC probes for mRNA, transcript/cDNA and genomic DNA. Recombinant host cells
CC expressing a native HK are useful for assaying compounds that stimulate
CC or inhibit HK function. Nucleic acids of the invention are also useful
CC for producing transgenic animals. Experimental data indicates that kinase
CC proteins of the present invention are expressed in humans in the eye
CC (retinoblastomas) and brain. The current sequence represents a rat amino
CC acid sequence related to to kinase protein of the invention, appearing as
CC SEQ ID 4 in the sequence listing, but is not referred to in the
XX specification
XX Sequence 505 AA;
XX
Query Match 93.7%; Score 2468; DB 5; Length 505;
Best Local Similarity 93.5%; Pred. No. 8.5e-221;
Matches 472; Conservative 11; Mismatches 22; Indels 0; Gaps 0;
Qy 1 MEGPVCVCCQDPRAELVERVAADIVTHLEADGGPEPTRNGVDPPPRARAASVIPGSTSR 60
Db 1 MEGPVCVCCQDPRAELVERVAADIVTHLEADGGPEPTRNGVDPPPRARAASVIPGSTSR 60
Qy 61 LLPARPSLSARKLSLOERPAGSYLEAQAGPYATGPASHISPRAWRPTIESHHVAISDAE 120
Db 61 PTPVRLPSLSARKLSLOERPAGSCLEAQVGFSTGPFASHMSPRAWRPTIESHHVAISDAE 120
Qy 121 DCVOLNQYKQSEIGKGYGVVRLAYNESEDRHYAMKVLSSKKLLKXQYGFPRPPRGSQ 180
Db 121 DCVOLNQYKQSEIGKGYGVVRLAYNEREDRHYAMKVLSSKKLLKXQYGFPRPPRGSQ 180
Qy 181 AAQGGPAKQLPLERVYQETAILKKLDHVNKKLIEVLDDPAEDNLKLVFLLRKGPMVE 240
Db 181 APQGGPAKQLPLERVYQETAILKKLDHVNKKLIEVLDDPAEDNLKLVFLLRKGPMVE 240
Qy 241 VPCDKPSEEARLYLRDVLGLYHCQKIVHRDIKPSNLLLGDDGHVKIADFGVSNQF 300
Db 241 VPCDKPPEEARLYLRDVLGLYHCQKIVHRDIKPSNLLLGDDGHVKIADFGVSNQF 300

QY 301 EGNDAQLSSTAGTAPAFMAPEAISDSQSGSKALDVWATGVTLYCFYVKGCPIDDFILA 360
 Db 301 EGNDAQLSSTAGTAPAFMAPEAISDTQSGSKALDVWATGVTLYCFYVKGCPIDDFILA 360
 QY 361 LHRKIKNEPVVFPPEEPISEELKDLILKMLDKNPETRIGVDPDIKLPWTKNGEELPSE 420
 Db 361 LHRKIKNEAVVFPPEEVESEELKDLILKMLDKNPETRIGVSDIKLPWTKNGEELPSE 420
 QY 421 EHCSSVVEVTEGEVKNVSLIPSWTTVILVKSMLKRKRSFGNPFEPQARREERSMSAPGNL 480
 Db 421 EHCSSVVEVTEGEVKNVSLIPSWTTVILVKSMLKRKRSFGNPFEPQARREERSMSAPGNL 480
 QY 481 LVKEGFGGKSPELPGVQDEAAS 505
 Db 481 LVKEGFGGKSPELPGVQDEAAS 505

RESULT 13

ADE61937

ID ADE61937 standard; protein; 505 AA.

XX

AC ADE61937;

DT 29-JAN-2004 (first entry)

XX

DE Rat Protein AAB46910, SEQ ID NO 7866.

XX

KW Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;

XX

KW chronic constriction injury; CCI; spared nerve injury; SNI; Chung.

XX

OS Rattus norvegicus.

XX

PN WO2003016475-A2.

XX

PD 27-FEB-2003.

XX

PF 14-AUG-2002; 2002WO-US025765.

XX

PR 14-AUG-2003; 2001US-0312147P.

XX

PR 01-NOV-2001; 2001US-0346382P.

XX

PR 26-NOV-2001; 2001US-0333347P.

XX

PA (GEO) GEN HOSPITAL CORP.

XX

PA (FAR) BAYER AG.

XX

PI Woolf C, D'urso D, Befort K, Costigan M;

XX

DR WPI; 2003-268312/26.

XX

DR GENBANK; AAB46910.

XX

PT New composition comprising two or more isolated polypeptides, useful for

XX

PT preparing a medicament for treating pain in an animal.

XX

PS Claim 1; Page; 1017pp; English.

XX

XX

CC The invention discloses a composition comprising two or more isolated rat

CC

CC or human polynucleotides or a polynucleotide which represents a fragment,

CC

CC derivative or allelic variation of the nucleic acid sequence. Also

CC

CC claimed are a vector comprising the novel polynucleotide, a host cell

CC

CC comprising the vector, a method for identifying a nucleotide sequence

CC

CC which is differentially regulated in an animal subjected to pain and a

CC

CC kit to perform the method, an array, a method for identifying an agent

CC

CC that increases or decreases the expression of the polynucleotide sequence

CC

CC that is differentially expressed in neuronal tissue of a first animal

CC

CC subjected to pain, a method for identifying a compound which regulates

CC

CC the expression of a polynucleotide sequence which is differentially

CC

CC expressed in an animal subjected to pain, a method for identifying a

CC

CC compound that regulates the activity of one or more of the

CC

CC polynucleotides, a method for producing a pharmaceutical composition, a

CC

CC method for identifying a compound or small molecule that regulates the

CC

CC activity in an animal of one or more of the polypeptides given in the

CC

CC specification, a method for identifying a compound useful in treating

CC

CC pain and a pharmaceutical composition comprising the one or more
 CC polypeptides or their antibodies. The polynucleotide or the compound that
 CC modulates its activity is useful for preparing a medicament for treating
 CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
 CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
 CC therapy). The sequence presented is a rat protein (shown in Table 2 of
 CC the specification) which is differentially expressed during pain. Note:
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic form directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.

XX

SQ Sequence 505 AA;

Query Match 93.7%; Score 2468; DB 7; Length 505;

Best Local Similarity 93.5%; Pred. No. 8.5e-22;

Matches 472; Conservative 11; Mismatches 22; Indels 0; Gaps 0;

QY 1 MEGGPAVCCODPRAELVERVAALDVTHLEADGGPEPTRNGVDPPPRARAASVIPGSTSR 60
 Db 1 MERSPAVCCODPRAELVERVAALDVTHLEADGGPEPTRNGVDPPPRARAASVIPGSASR 60
 QY 61 LLPAPSLSARKLSLQERPAQSYLEAQAGFYATGPASHISPRARRPTTIESHHVAISDAE 120
 Db 61 PTPVRPSLSARKPSLQERPAQSYLEAQAGFYATGPASHISPRARRPTTIESHHVAISDTE 120
 QY 121 DCVQLNQYKLOSEIGKAGYGVRLAYNESEDRHYAMKVLKSKLLKQYGFPRPPRGSQ 180
 Db 121 DCVQLNQYKLOSEIGKAGYGVRLAYNEREDRHYAMKVLKSKLLKQYGFPRPPRGSQ 180
 QY 181 AAGGPAKOLLPLERVYQEIATILKLDHNVVVKLIEVLDDPAEDNLVLFVLLRKGPMVE 240
 Db 181 APQGGPAKOLLPLERVYQEIATILKLDHNVVVKLIEVLDDPAEDNLVLFVLLRKGPMVE 240
 QY 241 VPCDKPFSEEQARLYLRDVLGLVYHCOKIVHRDIKPSNLLLGDDGHVKIADFGVSNQF 300
 Db 241 VPCDKPFSEEQARLYLRDVLGLVYHCOKIVHRDIKPSNLLLGDDGHVKIADFGVSNQF 300
 QY 301 EGNDAQLSSTAGTAPAFMAPEAISDSQSGSKALDVWATGVTLYCFYVKGCPIDDFILA 360
 Db 301 EGNDAQLSSTAGTAPAFMAPEAISDTQSGSKALDVWATGVTLYCFYVKGCPIDDFILA 360
 QY 361 LHRKIKNEPVVFPPEEPISEELKDLILKMLDKNPETRIGVDPDIKLPWTKNGEELPSE 420
 Db 361 LHRKIKNEAVVFPPEEVESEELKDLILKMLDKNPETRIGVSDIKLPWTKNGEELPSE 420
 QY 421 EHCSSVVEVTEGEVKNVSLIPSWTTVILVKSMLKRKRSFGNPFEPQARREERSMSAPGNL 480
 Db 421 EHCSSVVEVTEGEVKNVSLIPSWTTVILVKSMLKRKRSFGNPFEPQARREERSMSAPGNL 480
 QY 481 LVKEGFGGKSPELPGVQDEAAS 505
 Db 481 LVKEGFGGKSPELPGVQDEAAS 505

RESULT 14

ADE61941

ID ADE61941 standard; protein; 505 AA.

XX

AC ADE61941;

XX

DT 29-JAN-2004 (first entry)

XX

DE Rat Protein AAB46910, SEQ ID NO 7870.

XX

KW Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;

XX

KW chronic constriction injury; CCI; spared nerve injury; SNI; Chung.

XX

OS Rattus norvegicus.

XX

PN WO2003016475-A2.

XX

XX 27-FEB-2003.

XX

PF 14-AUG-2002; 2002WO-US025765.
XX 14-AUG-2001; 2001US-0312147P.
PR 01-NOV-2001; 2001US-0346382P.
PR 26-NOV-2001; 2001US-0333347P.
XX (GHEO) GEN HOSPITAL CORP.
FA (FARB) BAYER AG.
XX Woolf C, D'urso D, Befort K, Costigan M;
XX WPI; 2003-268312/26.
DR GENBANK; AAB46910.
XX
XX New composition comprising two or more isolated polypeptides, useful for
FT preparing a medicament for treating pain in an animal.
XX
XX Claim 1; Page; 1017pp; English.
XX
XX The invention discloses a composition comprising two or more isolated rat
CC or human polynucleotides or a polynucleotide which represents a fragment,
CC derivative or allelic variation of the nucleic acid sequence. Also
CC claimed are a vector comprising the novel polynucleotide, a host cell
CC comprising the vector, a method for identifying a nucleotide sequence
CC which is differentially regulated in an animal subjected to pain and a
CC kit to perform the method, an array, a method for identifying an agent
CC that increases or decreases the expression of the polynucleotide sequence
CC that is differentially expressed in neuronal tissue of a first animal
CC subjected to pain, a method for identifying a compound which regulates
CC the expression of a polynucleotide sequence which is differentially
CC expressed in an animal subjected to pain, a method for identifying a
CC compound that regulates the activity of one or more of the
CC polynucleotides, a method for producing a pharmaceutical composition, a
CC method for identifying a compound or small molecule that regulates the
CC activity in an animal of one or more of the polypeptides given in the
CC specification, a method for identifying a compound useful in treating
CC pain and a pharmaceutical composition comprising the one or more
CC polypeptides or their antibodies. The polynucleotide or the compound that
CC modulates its activity is useful for preparing a medicament for treating
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC therapy). The sequence presented is a rat protein (shown in Table 2 of
CC the specification) which is differentially expressed during pain. Note:
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 505 AA;

Query Match 93.7%; Score 2468; DB 7; Length 505;
Best Local Similarity 93.5%; Pred. No. 8.5e-221;
Matches 472; Conservative 11; Mismatches 22; Indels 0; Gaps 0;

Qy 1 MEGSPAVCCOPRAELVERVAIDVTHLEADGGPEPTENGVDPPRAASVPIGTSR 60
Db 1 MERSPAVCCOPRAELVERVAISVAHLEAEAGEPEASNGVDPPRAASVPIGTSR 60
Qy 61 LLPARPSLSARKLSIQERPAQSYLEAQAPYATGPASHISPAWRPRTIESHHVAISDAE 120
Db 61 PTPVRPSLSARKFSIQERPAQSCLAEQVGPYSTGPASHMSPAWRPRTIESHHVAISDTE 120
Qy 121 DCVQLNQYKLSQKYGAYGVVRLAYNESDRHYAMKVLKSKLLKQYGFPRPPRPGSQ 180
Db 121 DCVQLNQYKLSQKYGAYGVVRLAYNERDRHYAMKVLKSKLLKQYGFPRPPRPGSQ 180
Qy 181 AAGGPAKQLPLERVYQEIATLKLHDVNVVKLI EVLDDPAEDNLVLFDLIRKGPYME 240
Db 181 APQGGPAKQLPLERVYQEIATLKLHDVNVVKLI EVLDDPAEDNLVLFDLIRKGPYME 240
Qy 241 VPCDPPFEEQARLYRDLVILEYHLCQKIVHRDIKPSNLLGDDGHVKIADFGVSNQF 300
Db 241 VPCDPPFEEQARLYRDLVILEYHLCQKIVHRDIKPSNLLGDDGHVKIADFGVSNQF 300

Qy 301 EGNDQQLSSTAGTAPAFMAPEAISDSQSGFSKALDWNATGVTLYCFYVKGCPFIIDFFILA 360
Db 301 EGNDQQLSSTAGTAPAFMAPEAISDTGSGFSKALDWNATGVTLYCFYVKGCPFIIDFFILA 360
Qy 361 LHRKIKNEPVVPEPEISBELKDLILKMLDKNPETRIGVDPDIKHPWVTKNGEPLPSE 420
Db 361 LHRKIKNEAVVPEPEVSEBELKDLILKMLDKNPETRIGVSDIKLHPWVTKHGEPLPSE 420
Qy 421 EEHCSVVTEGEVKNVRLIPSWTTVLVKSMRLKRSFGNPFEPQARREERSMSAPGNL 480
Db 421 EEHCSVVVTEEEVKNSVKLIPSWTTVLVKSMRLKRSFGNPFEPQARREERSMSAPGNL 480
Qy 481 LVKSGFGEGGKSPELPGVQDEDAAS 505
Db 481 LLKEGCGEGGKSPELPGVQDEDAAS 505
RESULT 15
AA44239
ID AA44239 standard; protein; 540 AA.
XX
AC AA44239;
XX
DT 28-FEB-2000 (first entry)
XX Human cell signalling protein-2.
DE Cell signalling protein-2; CSIGP-2; cell proliferation; arteriosclerosis;
XX inflammatory disorder; cirrhosis; cancer; hepatitis; AIDS;
KW Addison's disease; multiple sclerosis.
XX Homo sapiens.
OS
XX
FH Key Location/Qualifiers
FT Modified-site 11 /note= "Potential phosphorylation site"
FT Modified-site 26 /note= "Potential phosphorylation site"
FT Modified-site 30 /note= "Potential phosphorylation site"
FT Modified-site 56 /note= "Potential phosphorylation site"
FT Modified-site 95 /note= "Potential phosphorylation site"
FT Modified-site 100 /note= "Potential phosphorylation site"
FT Modified-site 137 /note= "Potential phosphorylation site"
FT Modified-site 145 /note= "Potential phosphorylation site"
FT Region /note= "Potential phosphorylation site"
FT 165..446 /label= Signature_sequence
FT Modified-site 166 /note= "Potential phosphorylation site"
FT Modified-site 197 /note= "Potential phosphorylation site"
FT Modified-site 234 /note= "Potential phosphorylation site"
FT Modified-site 280 /note= "Potential phosphorylation site"
FT Modified-site 305 /note= "Potential phosphorylation site"
FT Modified-site 358 /note= "Potential phosphorylation site"
FT Modified-site 362 /note= "Potential phosphorylation site"
FT Modified-site 367 /note= "Potential phosphorylation site"
FT Modified-site 456 /note= "Potential phosphorylation site"
FT Modified-site 460 /note= "Potential phosphorylation site"
FT Modified-site 462 /note= "Potential glycosylation site"

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 25, 2004, 05:38:15 ; Search time 5991 Seconds
(without alignments)
7551.528 Million cell updates/sec

Title: US-10-690-617-1_COPY_173_1687

Perfect score: 1515

Sequence: 1 atggagggggtccagctgt.....aggaagcagggtgcattcc 1515

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

RST:*

1: em_estba:*

2: em_esthum:*

3: em_estin:*

4: em_estmu:*

5: em_estov:*

6: em_estpl:*

7: em_estro:*

8: em_htc:*

9: gb_est1:*

10: gb_est2:*

11: gb_htc:*

12: gb_est3:*

13: gb_est4:*

14: gb_est5:*

15: em_estfun:*

16: em_estom:*

17: em_gss_hum:*

18: em_gss_inv:*

19: em_gss_pln:*

20: em_gss_vrt:*

21: em_gss_fun:*

22: em_gss_mam:*

23: em_gss_mus:*

24: em_gss_pro:*

25: em_gss_rod:*

26: em_gss_phg:*

27: em_gss_vrl:*

28: gb_gss1:*

29: gb_gss2:*

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1508.6	99.6	1518	29 AY416154	AY416154 Homo sapi
2	1222.6	80.7	1518	29 AY416156	AY416156 Mus muscu
3	1053.8	70.0	1445	29 AY416155	AY416155 Pan trogl
4	940	62.0	1201	9 AL539375	AL539375 AL539375

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

5	892.4	58.9	1034	13	EX401350
6	889.6	58.7	1201	13	EX395129
7	779.4	51.4	1096	12	BM805189
8	777.2	51.3	939	13	BU552890
9	742.4	49.0	880	12	EG705317
10	732	48.3	910	13	BQ932168
11	723	47.7	1066	13	BQ277875
12	722	47.7	972	13	EX459685
13	692.6	45.7	893	12	EG396486
14	680.6	44.9	896	13	EX328136
15	668.6	44.1	784	13	BQ571613
16	658.4	43.5	685	12	EG701043
17	646.6	42.7	786	12	BI819446
18	639.8	42.2	2508	11	AK031399
19	633.2	41.8	2854	11	AK032070
20	589.8	38.9	3139	11	AK044660
21	570.6	37.7	727	12	BM950087
22	556	36.7	1127	13	BU902211
23	555.8	36.7	749	12	BI756153
24	529.2	34.9	648	12	EG294573
25	502.6	33.2	589	13	BU671280
26	468	30.9	636	12	BI546778
27	462.8	30.5	515	12	BM087446
28	439	29.0	900	12	BI254425
29	429.8	28.4	841	13	BU115704
30	417	27.5	496	14	CB712914
31	411.4	27.2	672	13	BY735290
32	398.6	26.3	1201	9	AL582916
33	397.6	26.2	414	12	BM150448
34	395.4	26.1	915	10	BE562166
35	389.2	25.7	857	13	BU420206
36	381.4	25.2	894	13	BU367227
37	373.4	24.6	438	13	EX488903
38	371	24.5	503	12	BM951274
39	370.6	24.5	727	12	BM952005
40	368.8	24.3	700	10	BE887687
41	366.6	24.2	750	12	BM718658
42	365.2	24.1	1120	14	CD498154
43	362	23.9	905	13	EX350723
44	361	23.8	703	28	AQ629789
45	351.6	23.2	881	13	EX328334

ALIGNMENTS

RESULT 1	AY416154	Homo sapiens CAMK1 gene, partial sequence, genomic survey sequence.	1518 bp	DNA linear	GSS 17-DEC-2003
LOCUS	AY416154	Genomic survey sequence.			
DEFINITION	AY416154.1	GI:39772114			
ACCESSION	AY416154				
VERSION	AY416154.1	GI:39772114			
KEYWORDS	GSS.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	1	(bases 1 to 1518)			
AUTHORS	Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarimal, A., Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B., Ferrisera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J., Adams, M.D. and Cargill, M.				
TITLE	Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios				
JOURNAL	Science	302 (5652),	1960-1963	(2003)	
PUBMED	14671302				
REFERENCE	2	(bases 1 to 1518)			
AUTHORS	Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarimal, A., Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B., Ferrisera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J., Adams, M.D. and Cargill, M.				
TITLE	Direct Submission				

```
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source 1..1518
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
<1..>1518
/gene="CAMKK1"
/locus_tag="HCM5796"
ORIGIN
Query Match 99.6%; Score 1508.6; DB 29; Length 1518;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1511; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 ATGAGGGGGTCCAGCTGCTGTGCTCCAGGATCCTCGGCAGAGCTGGTAGAACGGGTG 60
DB 1 ATGAGGGGGTCCAGCTGCTGTGCTCCAGATCCTCGGCAGAGCTGGTAGAACGGGTG 60
QY 61 GCAGCCATCGATGTAATCACTTGGAGGAGCGAGATGCTGGCCAGAGCCCTACTAGAAAC 120
DB 61 GCAGCCATCGATGTAATCACTTGGAGGAGCGAGATGCTGGCCAGAGCCCTACTAGAAAC 120
QY 121 GTGTGGAACCCCAACCAACGGGCGAGAGTGGCTCTGTGATCCCTGGCAGTACTTCAAGA 180
DB 121 GTGTGGAACCCCAACCAACGGGCGAGAGTGGCTCTGTGATCCCTGGCAGTACTTCAAGA 180
QY 181 CTGCTCCACCGCGCTAGCCTCTGACCCAGAGAGCTTCCCTACAGGAGCGGCCAGCA 240
DB 181 CTGCTCCACCGCGCTAGCCTCTGACCCAGAGAGCTTCCCTACAGGAGCGGCCAGCA 240
QY 241 GGAAGCTATCTGAGGCGCAGGCTGGCCCTTATGCCACGGGGCTGCGACCATCTCC 300
DB 241 GGAAGCTATCTGAGGCGCAGGCTGGCCCTTATGCCACGGGGCTGCGACCATCTCC 300
QY 301 CCCGGGCTGGGGAGGCCACCATGAGTCCACACAGTGGCCATCTCAGATGCAGAG 360
DB 301 CCCGGGCTGGGGAGGCCACCATGAGTCCACACAGTGGCCATCTCAGATGCAGAG 360
QY 361 GACTGGCTGAGTGAACAGTCAAGCTCAGAGTGCAGATGGCAAGGCTGCTACGGT 420
DB 361 GACTGGCTGAGTGAACAGTCAAGCTCAGAGTGCAGATGGCAAGGCTGCTACGGT 420
QY 421 GTGGTGAAGCTGCTTACAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 480
DB 421 GTGGTGAAGCTGCTTACAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 480
QY 481 AAAAGAGTACTGAAGAGTATGGCTTTCCAGTGCCTCCCGAGAGGTCCTCCAG 540
DB 481 AAAAGAGTACTGAAGAGTATGGCTTTCCAGTGCCTCCCGAGAGGTCCTCCAG 540
QY 541 GCTGCCAGGAGGACCAAGCAGTGTGCTGCCCTGAGGCGGTGTACCGAGATT 600
DB 541 GCTGCCAGGAGGACCAAGCAGTGTGCTGCCCTGAGGCGGTGTACCGAGATT 600
QY 601 GCATCTGAAGAGTGGACCAAGTGAATGTGGTCAAACTGATCGAGGTCTGGATGAC 660
DB 601 GCATCTGAAGAGTGGACCAAGTGAATGTGGTCAAACTGATCGAGGTCTGGATGAC 660
QY 661 CCAGCTGAGCAACCTCTATTGTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 720
DB 661 CCAGCTGAGCAACCTCTATTGTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 720
QY 721 GTGCCCTGTGACAAGCCCTTCTCGGAGGACCAAGTGTGCTTACCTCGGGAGCGTCATC 780
DB 721 GTGCCCTGTGACAAGCCCTTCTCGGAGGACCAAGTGTGCTTACCTCGGGAGCGTCATC 780
QY 781 CTGGGCTCGAGTACTTGCATCTGCAGAGATGCTGCAGAGGACATCAAGCCATCAAC 840
DB 781 CTGGGCTCGAGTACTTGCATCTGCAGAGATGCTGCAGAGGACATCAAGCCATCAAC 840
```

```
841 CTGCTCTCGGGGATGATGGGACGTCGTAAGATCGCCGACTTTTGGCGTCAGCAACGAGTTT 900
841 CTGCTCTCGGGGATGATGGGACGTCGTAAGATCGCCGACTTTTGGCGTCAGCAACGAGTTT 900
QY 901 GAGGGGAACGACGCTCAGCTGTCCAGCAGCGCGGACCCAGCAATTCATGGCCCCCGAG 960
901 GAGGGGAACGACGCTCAGCTGTCCAGCAGCGCGGACCCAGCAATTCATGGCCCCCGAG 960
QY 961 GCCATTTCTGATTCGGGCCAGAGCTTCAGTGGGAAGGCTTGGATGTATGGCCACTGGC 1020
961 GCCATTTCTGATTCGGGCCAGAGCTTCAGTGGGAAGGCTTGGATGTATGGCCACTGGC 1020
QY 1021 GTACAGTTGATCTGCTTTGCTATGGGAAGTGCCTATCATGACAGATTCATCTCTGGCC 1080
1021 GTACAGTTGATCTGCTTTGCTATGGGAAGTGCCTATCATGACAGATTCATCTCTGGCC 1080
QY 1081 CTCACAGGAAGATCAAGATGAGCCGCTGCTTCTCAGGAGCCAGAAATCAGCGAG 1140
1081 CTCACAGGAAGATCAAGATGAGCCGCTGCTTCTCAGGAGCCAGAAATCAGCGAG 1140
QY 1141 GAGCTCAAGGACCTGATCTCTGAAGATGTTAGACAAGAATCCCGAGACGAGAATTTGGGGTG 1200
1141 GAGCTCAAGGACCTGATCTCTGAAGATGTTAGACAAGAATCCCGAGACGAGAATTTGGGGTG 1200
QY 1201 CCAGACATCAAGTTCACCCCTTGGGTGACCAAGACGGGAGGAGCCCTTCTCTCGGAG 1260
1201 CCAGACATCAAGTTCACCCCTTGGGTGACCAAGACGGGAGGAGCCCTTCTCTCGGAG 1260
QY 1261 GAGGAGCACTGACGCTGGTGGAGGTGACAGAGGGGAGGTTAAGAACTCAGTCAGGCTC 1320
1261 GAGGAGCACTGACGCTGGTGGAGGTGACAGAGGGGAGGTTAAGAACTCAGTCAGGCTC 1320
QY 1321 ATCCGAGCTGACGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1380
1321 ATCCGAGCTGACGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1380
QY 1381 AACCCGTTTACGCCCCCAGCAGCGGAGGAGGATCCATGCTGCTCCAGGAAACCTA 1440
1381 AACCCGTTTACGCCCCCAGCAGCGGAGGAGGATCCATGCTGCTCCAGGAAACCTA 1440
QY 1441 CTGCTGAAAGAGGGTTTGGTGAAGGGGCAAGAGCCAGAGCTCCCGCGCTCCAGAA 1500
1441 CTGCTGAAAGAGGGTTTGGTGAAGGGGCAAGAGCCAGAGCTCCCGCGCTCCAGAA 1500
QY 1501 GACGAGCTCCATCC 1515
1501 GACGAGCTCCATCC 1515
RESULT 2
AY416156 1518 bp DNA linear GSS 17-DEC-2003
LOCUS Mus musculus CAMKK1 gene, VIRTUAL TRANSCRIPT, partial sequence,
DEFINITION genomic survey sequence.
ACCESSION AY416156
VERSION AY416156.1 GI:39772116
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 1518)
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sainsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 1518)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
```

Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.

TITLE Direct Submission

JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA

COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment.

FEATURES Location/Qualifiers

source

1..1518

/organism="Mus musculus"

/mol_type="genomic DNA"

/db_xref="taxon:10090"

<1..>1518

/gene="CAMK1"

/locus_tag="HCM5796"

ORIGIN

Query Match 80.7%; Score 1222.6; DB 29; Length 1518;

Best Local Similarity 87.4%; Pred. No. 9.7e-260;

Matches 1324; Conservative 0; Mismatches 191; Indels 0; Gaps 0;

```
QY 1 ATGGAGGGGTCCAGCTGTCTGCTGCCAGATCCTCGGCGAGAGCTGGTAGAAGGGTG 60
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1 ATGGAGAGTGGCCCGAGCGCTGTCTGCCAAGACCCCTGTGCAGAGCTGGTAGAAGGGTG 60
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 61 GCAGCCATCGATGTGACTCACTTGGAGAGGCGAGATGGTGGCCAGAGCCCTACTAGAAAC 120
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 61 GCGGCCATCAATGTGGCCCACTTGGAGAAGCAGATGAGGGTCCAGAGCCCTGCCAGGAT 120
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 121 GGTGTGACCCCCCAACACCGGCCAGAGCTGCCTCTGTGATCCCTGGCAGTACTTCAAGA 180
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 121 GGTGTGATCTCCACCCCGGCCAGAGCTGCCTCCGTGATCCCTGGCAGTACTTCAAGA 180
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 181 CTGTCTCCAGCCCGGCTAGCTCTACCCAGGAGCTTCCCTACAGAGCGGCCAGCA 240
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 181 CCCACTCCAGTGGCCCGCCAGCTCTCGGCTGAGAAAGTTCTCCCTGCAGGAACGACGCT 240
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 241 GGAAGCTATCTGGAGGCGCAGGCTGGGCTTATGCCAGGGGCGCTGCCAGGCCACATCTCC 300
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 241 GGAAGCTCTTGGGGCTCAAGTTGGGCTTACTCTACAGAGCTGCCAGTCAATCTCT 300
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 301 CCCGGGCTCGGAGAGCCCAACATCGAGTCCCAACACGTTGGCATCTCAGATGCAGAG 360
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 301 CCTCGGCTCTGGCGAGAGCCCAACATCGAGTCCCAACCGTGTGGCATCTCAGACACAGAG 360
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 361 GACTGCGTGCAGTGAACCAAGTACAAAGCTGCAGAGTGAGATTGGCAAGGTCCTACGGT 420
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 361 GACTGTGTCAACTGAACCAAGTACAAAGCTGCAGAGTGAGATTGGCAAGGTCCTATGGT 420
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 421 GTGTGTAGGCTGGCTACAACGAAAGTGAAGACAGACACTATGCAATGAAAGTCTTTTCC 480
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 421 GTGTGTAGGCTGGCTACAACGAAAGTGAAGACAGACACTATGCAATGAAAGTCTTTTCC 480
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 481 AAAAAGAGTTACTGAAGCAGTATGGCTTCCACGTCGCCCTCCCCGAGAGGGTCCAG 540
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 481 AAAAAGAGTTACTGAAGCAGTATGGCTTCCACGTCGCCCTCCCCGAGAGGGTCCCAA 540
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 541 GCTGCCAGGAGGACCAAGCAAGCTGTGCCCTGGAGCGGCTACACGAGATT 600
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 541 GCTACCCAGGAGGCGCAAGCAAGCTGTGCCCTGGAGCGGCTGTATCAGAGATT 600
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 601 GCCATCTTGAAGAGCTGGACCACTGTAATGTGGTCAAACTGATCGAGTCTCTGATGAC 660
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 601 GCCATCTTGAAGAGCTGGACCACTGTAATGTAGTCAAACTGATCGAGTCTCTGATGAT 660
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 661 CCAGTGTAGGACAACTCTATTGTGTGTGACCTCTCTGAGAAAGGGCCCGCTCATGAA 720
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 661 CCCGCTGAAGACAACTCTATTGTGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 720
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 721 GTGCCCTGTGACAGCCCTTCTCGAGAGGACGAGCTCGCCTTACTCTCGGAGCGCTCATC 780
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 721 GTGCCCTGTGACAGCCCTTCTCGAGAGGACGAGCTCGCCTTACTCTCGGAGCATCATC 780
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```


BP 191 91006 EVRY cedex - France
 Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr
 Library was constructed by Life Technologies, a division of
 Invitrogen. This sequence belongs to sequence cluster 8786.r For
 more information about this cluster, see
 http://www.genoscope.cns.fr/
 cgi-bin/cluster.cgi?seq=CS0DF033AH09QF1&cluster=8786.r. Contact :
 Feng Liang Email : fliang@lifetech.com URL :
 http://fulllength.invitrogen.com/ InvitroGen Corporation 1600
 Faraday Avenue Genoscope sequence ID : CS0DF033AH09QP1.
 Location/Qualifiers
 1. .1201
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0DF033Y017"
 /tissue_type="FETAL BRAIN"
 /dev_stage="fetal"
 /clone_lib="Homo sapiens FETAL BRAIN"
 /note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
 was primed with a NotI-oligo(dT) primer. Five prime end
 enriched, double-strand cDNA was digested with Not I and
 cloned into the Not I and EcoRV sites of the pCMVSPORT 6
 vector. Library was not normalized."
 ORIGIN
 Query Match 62.0%; Score 940; DB 9; Length 1201;
 Best Local Similarity 97.2%; Pred. No. 3 3e-197;
 Matches 1005; Conservative 5; Mismatches 16; Indels 8; Gaps 5;
 QY 6 GGGGGTCCAGCTGTCTGCTGCCAGATCCTCGGCAGAGCTGGTAGAACGGGTGGCAGC 65
 Db 65 GGGGGTCCAGCTGTCTGCTGCCAGATCCTCGGCAGAGCTGGTAGAACGGGTGGCAGC 124
 QY 66 CATCATGTGACTCACTTGGAGGAGCAGATGGTGGCCAGAGCCTACTAGAAACGGTGT 125
 Db 125 CATCATGTGACTCACTTGGAGGAGCAGATGGTGGCCAGAGCCTACTAGAAACGGTGT 184
 QY 126 GGACCCCCCACCAGCGGCAGAGCTGCCCTCTGTATCCCTGGCAGTACTCAAGACTGCT 185
 Db 185 GGACCCCCCACCAGCGGCAGAGCTGCCCTCTGTATCCCTGGCAGTACTCAAGACTGCT 244
 QY 186 CCAGCCGGCGCTAGCCTCTCAGCCAGGAAGCTTTCCCTACAGGAGCGCCAGCAGGAAG 245
 Db 245 CCAGCCGGCGCTAGCCTCTCAGCCAGGAAGCTTTCCCTACAGGAGCGCCAGCAGGAAG 304
 QY 246 CTATCTGGAGGCGAGGCTGGGCTTATGCCAGGGGCTTCCAGGCCATCTCCCGCCG 305
 Db 305 CTATCTGGAGGCGAGGCTGGGCTTATGCCAGGGGCTTCCAGGCCATCTCCCGCCG 364
 QY 306 GGCCTGGCGAGGCGCCACCACCTGAGTCCACACACGTTGGCCATCTCAGATGCAGAGACTG 365
 Db 365 GGCCTGGCGAGGCGCCACCACCTGAGTCCACACACGTTGGCCATCTCAGATGCAGAGACTG 424
 QY 366 CTGTCAGCTGAACCCAGTCAAGCTGCAGAGTGAAGTTCGAGGAGGCTTACGGTGTGT 425
 Db 425 CTGTCAGCTGAACCCAGTCAAGCTGCAGAGTGAAGTTCGAGGAGGCTTACGGTGTGT 484
 QY 426 GAGGCTGGCTACACGAAGTGAACACAGACACTATGCAATGAAGTCTTCCAAA 485
 Db 485 GAGGCTGGCTACACGAAGTGAACACAGACACTATGCAATGAAGTCTTCCAAA 544
 QY 486 GAAGTTTACTGAAGCAGTATGGCTTTCCAGTCCGCTCCCGCAGAGGGTCCCGAGCTGC 545
 Db 545 GAAGTTTACTGAAGCAGTATGGCTTTCCAGTCCGCTCCCGCAGAGGGTCCCGAGCTGC 604
 QY 546 CCAGGAGGACCCAGCAGCAGCTGCTGCCCTTGGAGCGGGTGTACACAGAGATTGCCAT 605
 Db 605 CCAGGAGGACCCAGCAGCAGCTGCTGCCCTTGGAGCGGGTGTACACAGAGATTGCCAT 664
 QY 606 CTTGAAGAGCTGGACCAAGTGAATGGTCAACATGATCGAGTCTCGATGACCCAGC 665
 Db 665 CTTGAAGAGCTGGACCAAGTGAATGGTCAACATGATCGAGTCTCGATGACCCAGC 724

QY 666 TGAGACAACTCTATTTGGTGTGGTCTCTCTGAGAAAGGGCCCGTCATGGAAGTGC 725
 Db 725 TGAGACAACTCTATTTGGTGTGGTCTCTCTGAGAAAGGGCCCGTCATGGAAGTGC 784
 QY 726 CTGTGACAAAGCCCTTCTCGGAGGAGCAAGCTGCTCTACCTGCGGAGCTCATCTGGG 785
 Db 785 CTGTGACAAAGCCCTTCTCGGAGGAGCAAGCTGCTCTACCTGCGGAGCTCATCTGGG 844
 QY 786 COTCGAGTACTTGCACTGCCAGAGATCGTCCACAGGACATCAAGCCATCAACCTGCT 845
 Db 845 COTCGAGTACTTGCACTGCCAGAGATCGTCCACAGGACATCAAGCCATCAACCTGCT 904
 QY 846 COTGGGGGATGATGGGCACGTGAAGATCCCGACTTTGGCGTCAAGCAACCACTTTGA 904
 Db 905 COTGGGGGATGATGGGCACGTGAAGATCCCGACTTTGGCGTCAAGCAACCACTTTGA 964
 QY 905 GGAACGAGCTCAGCTGTCCAGCAGCGGGGAGAACCCAGCATTCTATGGCCCCGAGGCA 964
 Db 965 GGAACGAGCTCAGCTGTCCAGCAGCGGGGAGAACCCAGCATTCTATGGCCCCC 1021
 QY 965 TTTCGATTTCGGGCCAGAGCTTCACTGGGAAGCCTTGGATGTATGGGCCACTGGCGTCA 1024
 Db 1022 TTTCGATTTCGGGCC--AGAGCTCAGTGGGAAGSCTTGGATGTATGGGC--ACTGGGGA 1077
 QY 1025 COTTCCTACTGTTT 1038
 Db 1078 COTTCCTACTGHTTK 1091
 RESULT 5
 BX401350
 LOCUS
 DEFINITION
 BX401350 Homo sapiens HELA CELLS COT 25-NORMALIZED Homo sapiens
 CDNA clone CS0DK012YG22 5-PRIME, mRNA sequence.
 ACCESSION
 BX401350
 VERSION
 BX401350.1 GI:30630361
 KEYWORDS
 EST.
 SOURCE
 Homo sapiens (human)
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE
 1 (bases 1 to 1034)
 AUTHORS
 Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
 TITLE
 Full-length cDNA libraries and normalization
 JOURNAL
 Unpublished (2001)
 COMMENT
 Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 EVRY cedex - France
 Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr
 Library was constructed by Life Technologies, a division of
 Invitrogen. This sequence belongs to sequence cluster 8786.r For
 more information about this cluster, see
 http://www.genoscope.cns.fr/
 cgi-bin/cluster.cgi?seq=CS0DK012BD11QP1&cluster=8786.r. Contact :
 Feng Liang Email : fliang@lifetech.com URL :
 http://fulllength.invitrogen.com/ InvitroGen Corporation 1600
 Faraday Avenue Genoscope sequence ID : CS0DK012BD11QP1.
 Location/Qualifiers
 1. .1034
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0DK012YG22"
 /cell_type="HELA"
 /cell_line="HELA"
 /clone_lib="HELA CELLS COT 25-NORMALIZED"
 /note="1st strand cDNA was primed with a NotI-oligo(dT)
 primer. Five prime end enriched, double-strand cDNA was
 digested with Not I and cloned into the Not I and EcoRV
 sites of the pCMVSPORT 6 vector. Library was normalized."
 ORIGIN
 source

cloned (EcorV site is destroyed upon cloning). Average insert size 1.4 kb, insert size range 0.9-4 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 012."

Tracking Code 022									
ORIGIN									
Query Match	51.4%; Score 779.4; DB 12; Length 1096;								
Best Local Similarity	93.1%; Pred. No. 1.2e-161;								
Matches	905; Conservative 0; Mismatches 52; Indels 15; Gaps 8;								
Qy	1	ATGAGGGGGTCCAGCTGTCGTCTGCGCAGATCTCTCGGCAGAGTGTGTAGAACGGGTG	60						
Db	78	ATGAGGGGGGTCCAGTGTCTGCTCCAGATCTCTCGGCAGAGCTGTGTAGAACGGGTG	137						
Qy	61	GCAGCCATCGATGTGACTCATTTGGAGGAGCAGATGTGTGCCAGAGCCTACTAGAAAC	120						
Db	138	GCAGCCATCGATGTGACTCATTTGGAGGAGCAGATGTTGCCCCAGAGCCTACTAGAAAC	197						
Qy	121	GGTGTGAGCCCCCACCACGGGCAGAGCTGCTCTGTGTATCTCTGGCAGTACTTCAAGA	180						
Db	198	GGTGTGAGCCCCCACCACGGGCAGAGTGCCTCTGTGTATCTCTGGCAGTACTTCAAGA	257						
Qy	181	CTGTCTCCAGCCCGGCTAGCCTCTCAGCCAGGAAGCTTTCCCTACAGGAGCGGCCAGCA	240						
Db	258	CTGCTCCCAGCCGGCTAGCCTCTCAGCCAGGAAGCTTTCCCTACAGGAGCGGCCAGCA	317						
Qy	241	GGAAGCTATCTGAGGGCGAGGCTGGGCCCTTATGCCACGGGGCTGCCAGCCACATCTCC	300						
Db	318	GGAAGCTATCTGAGGGCGAGGCTGGGCCCTTATGCCACGGGGCTGCCAGCCACATCTCC	377						
Qy	301	CCCCGGGCTTGGCGAGGGCCACCATCGAGTCCCACCACGCTGGGCCATCTCAGATGTCGAG	360						
Db	378	CCCCGGGCTTGGCGAGGGCCACCATCGAGTCCCACCACGCTGGGCCATCTCAGATGTCGAG	437						
Qy	361	GACTGCGTCAGGTGAACGAGTACAGCTGCAGAGTGAGATTGGCAAGGGTGCCTAGCGT	420						
Db	438	GACTGCGTCAGGTGAACCGAGTCAAGCTGCAGAGTGAGATTGGCAAGGGTGCCTAGCGT	497						
Qy	421	GTGTGAGGCTGGCCTAACAGAAAGTGAAGACAGACACTATGCAATGAAGTCTCTTCC	480						
Db	498	GTGTGAGGCTGGCCTAACAGAAAGTGAAGACAGACACTATGCAATGAAGTCTCTTCC	557						
Qy	481	AAAAAGAGTTACTGAAGCAGTATGGCTTTCCAGCTGCGCCTCCCCGAGAGGGTCCCGAG	540						
Db	558	AAAAAGAGTTACTGAAGCAGTATGGCTTTCCAGCTGCGCCTCCCCGAGAGGGTCCCGAG	617						
Qy	541	GCTGCCAGGAGACAGCCACGACCTGCTG-CCCTGTGAGCGGGTGTACACGAGGAT	599						
Db	618	GCTGCCAGGAGACAGCCACGACCTGCTGCCCTGTGAGCGGGTGTACACGAGGAT	677						
Qy	600	TGCCATCTGAAGAAGCTGGACCACGTGAATGTGGTCAAACTGATCGAGGTCCTGGATGA	659						
Db	678	TGCCATCTGAAGAAGCTGGACCACGTGAATGTGGTCAAACTGATCGAGGTCCTGGATGA	737						
Qy	660	CCGAGCTGAGGACAACTCTATTGG-TGTGTACCTCTCTCAGAAAGGG-CCCGTCAATG	717						
Db	738	CCGAGCTGAGGACCACTCTATTGGTGTGTACCTCTCTCAGAAAGGGCCCCGTCAATG	797						
Qy	718	GAACTGCCCTGTGACAA-GCCCTTCTCGGAGGAGCAAGCTGCGCCTCTACCTGCGGGAGCT	776						
Db	798	GAACTGCCCTGTGACAAAGGCCCTTCTCGGAGGAGCAAGCTGCGCCTCTACCTGCGGGAGCT	857						
Qy	777	CATCC-TGGGCTCGAGTACTTGACCTGCCAGAGATGCTCCACAGGGACATCAAGCCAT	835						
Db	858	CATCCTTTGGGCTCGAGTACTTGACCTGGCCAGAGATCGTTTCAAGGGACATTAAGCCAT	917						
Qy	836	CCAACTTGCTCCTGGGG----GATGATGGGCACGTGAAGATGCCCACTTTGGCGTTCAG	890						
Db	918	TCCAACTTGGTTCCTTGGGGGGATGATGGGCACCTGAAAGATCGGCCCACTTTGGGGGCA	977						
Qy	891	CAACC--AGTTTGAAGGGAACGACGCTCAGTGTTC---AGCAGGGGGGAACCCGACCA	945						

481	QY	AAAAAGAGTTACTGAACAGTATGGCTTTCCACGTGCGCCCTCCCCCGAGAGGGTCCCGAG	540
563	Db	AAAAAGAGTTACTGAACAGTATGGCTTTCCACGTGCGCCCTCCCCCGAGAGGGTCCCGAG	622
541	QY	GCTGCCACGGGAGGACCAGGCAACAGCTGCTGCCCTCGGAGCGGGGTGTACCAAGAGATT	600
623	Db	GTTGCCACGGGAGGACCAGCACCAGCAGCTGCTGCCCTCGGAGCGGGTGTACCAAGAGATT	682
601	QY	GCCATCTTGAAGAAGCTGGACCAACGTGAATGTGTCAAACCTGATCGAGGTCTTGGATGAC	660
683	Db	GCCATCTTGAAGAAGCTGGACCAACGTGAATGTGTCAAACCTGATCGAGGTCTTGGATGAC	742
661	QY	CCAGCTGAGGACAACTCTATTGGTGTTCGACCTCTGACAAGGGGCCGTCATGGAA	720
743	Db	CCAGCTGAGGACAACTCTATTGGTGTTCGACCTCTGACAAGGGGCCGTCATGGAA	802
721	QY	GTGCCCTGTGACAAGCCCTTCTCGGAGGAGCAAGCTGCGCTTACCTTCGCGGACGTCATC	780
803	Db	GTGCCCTGTGACAAGCCCTTCTCGGAGGAGCAAGCTGCGCTTACCTTCGCGGACGTCATC	862
781	QY	CTGGGCTTCGAGTACTTGGACATGCGCAGAAGATCGTCCAAGGGACATCAAGCCATCAAC	840
863	Db	CTGGGCTTCGAGTACTTGGACATGCGCAGAAGATCGTCCAAGGGACATCAAGCCATCAAC	922
841	QY	CTGCTCTCTGGGGGATGATGGGCACGTGAAGATCGCCGACTTTGGCGTCAGCAACCAGTTT	900
923	Db	CTGCTCTCTGGGGGATGATGGGCACGTGAAGATCGCCGACTTTGGCGTCA-SAACCAGTTT	981
901	QY	GAGGGGAA	908
982	Db	GAGGGGAA	989

RESULT	7
BM805189	
LOCUS	BM805189 1096 bp linear mRNA EST 05-MAR-2002
DEFINITION	AGENCOURT_6499735 NIH_MGC_124 Homo sapiens cDNA clone IMAGE:5728398 5'.. mRNA sequence.

ACCESSION	BM805189	
VERSION	BM805189.1	GI:19122012
KEYWORDS	EST.	
SOURCE	Homo sapiens	
ORGANISM	Homo sapiens (human)	
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
	S, mRNA sequence.	

REFERENCE
1 (bases 1 to 1096)
AUTHORS
NIH-MGC <http://mgc.nci.nih.gov/>
TITLE
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL
Unpublished (1999)
Mammalia, Euderia, Firmacea, Cecarimnia, Homidae, Homot.

CONTACT: Robert Strausberg, Ph.D.
Email: gsabbs@mail.nih.gov
Tissue Procurement: Invitrogen
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:

found through the I.M.A.G.E. Consortium
<http://image.llnl.gov>
 Plate: LLAMI2724 row: a column: 07
 High quality sequence stop: 633.

FEATURES	SOURCE
1. .1096	Location/Qualifiers
/organism="Homo sapiens"	
/mol_type="mRNA"	
/db_xref="taxon:9606"	
/clone="IMAGE:5728398"	
/tissue_type="hippocampus"	
/lab_host="DH10B"	
/clone_lib="NH_MGC_124"	
/note="Organ: brain; Vector (destroyed); Site: 2; Not; age 27. Library is oligo-d"	

Db 978 CAACCCANTTTAGGGGAACACACGCTCAACTGTTCCAGACGGGGGAACCCCAACCA 1037

QY 946 TTTCATGGCCCC 957

Db 1038 TTTCGGGCCCC 1049

RESULT 8

LOCUS BU552890

DEFINITION AGENCOURT_10373297 NIH_MGC_109 Homo sapiens cDNA clone

IMAGE:6576737 5', mRNA sequence.

ACCESSION BU552890

VERSION BU552890.1 GI:22903162

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS 1 (bases 1 to 939)

TITLE NIH-MGC http://mgi.nci.nih.gov/.

JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)

COMMENT Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgabbs-remail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: Rubin Laboratory

DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Plate: LICM2775 row: 1 column: 17

High quality sequence stop: 658.

Location/Qualifiers

1..939

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:6576737"

/tissue_type="teratocarcinoma, cell line"

/lab_host="DH10B (phage-resistant)"

/clone_lib="NIH_MGC_109"

/note="Organ: Ovary; Vector: pOTB7; Site 1: EcoRI; Site 2: XhoI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

ORIGIN

Query Match 51.3%; Score 777.2; DB 13; Length 939;

Best Local Similarity 98.7%; Pred. No. 3.3e-161;

Matches 815; Conservative 0; Mismatches 8; Indels 3; Gaps 3;

QY 685 GTGTTTGACCTCTGTGAAAGGGGCGCGTCATGGAAGTGCCCTGTGACAGCCCTTCTCG 744

Db 38 GTGTTTGACCTCTGTGAAAGGGGCGCGTCATGGAAGTGCCCTGTGACAGCCCTTCTCG 97

QY 745 GAGGACCAAGCTCGCTCTACTTGGCGGAGCGTCATCTGGGCTCGAGTACTTGCAGTGC 804

Db 98 GAGGACCAAGCTCGCTCTACTTGGCGGAGCGTCATGGAAGTGCCCTGTGACAGCCCTTCTCG 97

QY 805 CAGAAGATCGTCCACAGGACATCAAGCCATCAACCTGCTCTGGGAGTGTGGGAGTGTGGGAC 864

Db 158 CAGAAGATCGTCCACAGGACATCAAGCCATCAACCTGCTCTGGGAGTGTGGGAGTGTGGGAC 217

QY 865 GTGAAGATCGCGACTTTGGCGGTGACAAACAGTTTGGGGGAAACGACGCTCAGCTGTCC 924

Db 218 GTGAAGATCGCGACTTTGGCGGTGACAAACAGTTTGGGGGAAACGACGCTCAGCTGTCC 277

QY 925 AGCAGCGGGGAACCCAGCATTTCATGCCCCCGAGGGCCATTTCGATTCCGGCCAGAGC 984

Db 278 AGCAGCGGGGAACCCAGCATTTCATGCCCCCGAGGGCCATTTCGATTCCGGCCAGAGC 337

QY 985 TTCAGTGGAAAGGCTTCGATGTATGGGCCACTGCGCTCAGCTTCTACTGCTTCTCTAT 1044

Db 338 TTCAGTGGAAAGGCTTCGATGTATGGGCCACTGCGCTCAGCTTCTACTGCTTCTCTAT 397

QY 1045 GGGAAAGTCCCATTCATCGACGATTTCATCTCGCCCTCCACAGGAAGATCAAGAATGAG 1104

Db 398 GGGAAAGTCCCGTTTCATCGACGATTTCATCTCGCCCTCCACAGGAAGATCAAGAATGAG 457

QY 1105 CCCGTGCTGTTTTCCTGAGGAGCCAGAAATCAGCAGGAGCTCAAGGACCTGATCTCTGAAG 1164

Db 458 CCCGTGCTGTTTTCCTGAGGAGCCAGAAATCAGCAGGAGCTCAAGGACCTGATCTCTGAAG 517

QY 1165 ATGTTAGCAAGAATCCCGAGACGAGAAATGGGGTGCCAGACATCAAGTTCCACCCCTGG 1224

Db 518 ATGTTAGCAAGAATCCCGAGACGAGAAATGGGGTGCCAGACATCAAGTTCCACCCCTGG 577

QY 1225 GTGACCAAGAACGGGGAGGAGCCCTTCCTTCGGAGGAGGACACTGCAGCGTGGTGGAG 1284

Db 578 GTGACCAAGAACGGGGAGGAGCCCTTCCTTCGGAGGAGGACACTGCAGCGTGGTGGAG 637

QY 1285 GTGACAGAGGGGAGGTTAAGAACTCAGTCAGGCTATCCCGAGTGCACACCGTGAATC 1344

Db 638 GTGACAGAGGAGGAGGTTAAGAACTCAGTCAGGCTATCCCGAGTGCACACCGTGAATC 697

QY 1345 CTGCTGAAGTCCATGCTGAGGAGCGTTCCCTTTGGGAACCCGTTTGAGCCGCCAGGACGG 1404

Db 698 CTGCTGAAGTCCATGCTGAGGAGCGTTCCCTTTGGGAACCCGTTTGAGCCGCCAGGACGG 757

QY 1405 AGGGAAGAGCGATCCATGCTGCTCCAGGAACCTA-CTGGTGAAGAAGGG-TTGGTG 1462

Db 758 AGGGAAGAGCGATCCATGCTGCTCCAGGAACCTA-CTGGTGAAGAAGGG-TTGGTG 817

QY 1463 AAGGGGCAAGAGCCAGAGCTCCCGGC-GTCCAGGAAGAGGAGG 1507

Db 818 AAGGGGCAAGAGCCCAAGCTCCCGCGCGTCCAGGAAGACCAAG 863

RESULT 9

LOCUS BG705317

DEFINITION BG705317 NIH_MGC_95 Homo sapiens cDNA clone IMAGE:4820238 5', mRNA sequence.

ACCESSION BG705317

VERSION BG705317.1 GI:13979532

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS 1 (bases 1 to 880)

TITLE NIH-MGC http://mgi.nci.nih.gov/.

JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)

COMMENT Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgabbs-remail.nih.gov

Tissue Procurement: Miklos Palkovits, M.D., Ph.D.

cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki Toshiyuki and Piero Carninci (RIKEN)

DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Plate: LLAM10725 row: a column: 07

High quality sequence stop: 769.

Location/Qualifiers

1..880

/organism="Homo sapiens"

/mol_type="mRNA"

FEATURES

source

```
/db_xref="taxon:9606"
/clone="IMAGE:4820238"
/tissue_type="hippocampus"
/lab_host="DH10B"
/clone_lib="NIH_MGC_95"
/Note="Organ: Brain; Vector: pBluescriptR (modified
pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI
(gtcgag); Oligo-dT primed using primer
5'-TTTTTTTTTTTTTNN-3', size-selected for average
insert size 2.5 kb and normalized to R0T 5. This is a
primary library enriched for full-length clones and
constructed using the cap-trapper method (Carninci, in
preparation). Library constructed by M. Brownstein
(NIH/NHGRI, National Institutes of Health). Note: this
is a NIH_MGC Library."

ORIGIN
Query Match 49.0%; Score 742.4; DB 12; Length 880;
Best Local Similarity 95.7%; Pred. No. 1.6e-153;
Matches 827; Conservative 0; Mismatches 31; Indels 6; Gaps 6;

QY 26 GCCAGATCCTCGGCAGAGCTGTAGAACGGGTGCAGCCATCGATGACTCACTTGG 85
DB 4 GGCAGATCCTCGGCAGAGCTGTAGAACGGGTGCAGCCATCGATGACTCACTTGG 63

QY 86 AGGAGGCAGATGCTGCCAGAGCTTACTAGAAACGGTGTGACACCCGCCACACCGGCCA 145
DB 64 AGGAGGCAGATGCTGCCAGAGCTTACTAGAAACGGTGTGACACCCGCCACACCGGCCA 123

QY 146 GAGTGCCTCTGTGATCCCTGCGAGTACTAAGACTGTCCAGCCCGGCTAGCCTT 205
DB 124 GAGTGCCTCTGTGATCCCTGCGAGTACTAAGACTGTCCAGCCCGGCTAGCCTT 183

QY 206 CAGCAGGAGCTTCCCTACAGGCGGCCAGCAGGAGCTATCTGGAGGCGCAGGCTG 265
DB 184 CAGCAGGAGCTTCCCTACAGGCGGCCAGCAGGAGCTATCTGGAGGCGCAGGCTG 243

QY 266 GGCCTTATGCCACGGGCTGCCAGACATCTCCCGCGGCGCTGGCGAGGCCACCA 325
DB 244 GGCCTTATGCCACGGGCTGCCAGACATCTCCCGCGGCGCTGGCGAGGCCACCA 303

QY 326 TCGAGTCCACACGTCGTCATCTAGATCGAGAGACTGCGTGTGAGCTGAACGATCA 385
DB 304 TCGAGTCCACACGTCGTCATCTAGATCGAGAGACTGCGTGTGAGCTGAACGATCA 363

QY 386 AGCTGCAGAGTGAAGTTCGCAAGGGTCCCTACGGTGTGAGCTGGCTCAACGAAA 445
DB 364 AGCTGCAGAGTGAAGTTCGCAAGGGTCCCTACGGTGTGAGCTGGCTCAACGAAA 423

QY 446 GTGAGACAGACACTATGCAATGAAAGTCTTTTCCAAAAGAGTTACTGAAGCAGTATG 505
DB 424 GTGAGACAGACACTATGCAATGAAAGTCTTTTCCAAAAGAGTTACTGAAGCAGTATG 483

QY 506 GCTTTCCAGTGCCTCTCCCGAGAGGGTCCAGAGTCCCGAGGAGCAGCAGCAGCAGC 565
DB 484 GCTTTCCAGTGCCTCTCCCGAGAGGGTCCAGAGTCCCGAGGAGCAGCAGCAGCAGC 543

QY 566 AGCTGCTGCCCTGGAGCGGCTGTACAGGAGATTGCCATCTGTGAGAGAGCTGGACACG 625
DB 544 AGCTGCTGCCCTGGAGCGGCTGTACAGGAGATTG-CATCTGAAAGAGCTGGACACG 602

QY 626 TGAATGTGTCAAACTGATCAGGTCTGTGATGACCCAGCTGAGGACAACTCTATTGG 685
DB 603 TGAATGTGTCAAACTGATCAGGT-CTGATGACCCAGCTGAGGACAACTCTATTGG 661

QY 686 TGTTTGACCTCTGAGAAAGGCGCGTATGAAAGTCCCTGTGACAAAGCCCTTCTGG 745
DB 662 TG-TTGACCTCTGAGAAAGGCGG-CGCTCATGAAAGTCCCTGTGACAAAGCCCTTCTGG 719

QY 746 AGGAGCAGCTCGCTCTACTCGGGACGTCATCTCGGCGCTCGAGTACTGCACTGCC 805
DB 720 AGGAGCAGCTCGCTCTACTCGGGACGTCATCTCGGCGCTCGAGTACTGCACTGCC 779
```

```
806 AGAA-GATCGTCCACA-GGCATCAAGCCATCCAACTGCTCTCTGGGGATGATGGGCA 863
780 AGAAGATCGTCCACAGGGGACCTCAGCGCATCAACCTGCTCTCTGGGGGATGATGGCC 839
864 CGTGAGATCGCCGACTTTGGCGT 887
840 GGTAAAGCATCGGCAAGTAGGCGGT 863

BQ932168 910 bp mRNA linear EST 21-AUG-2002
AGENCOURT 8805197 Lupski_sciatic_nerve Homo sapiens cDNA clone
IMAGE:6196986 5', mRNA sequence.
BQ932168
BQ932168.1 GI:22347199
EST.
Homo sapiens (human)
Homo sapiens
Homo sapiens
Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NIH-MGC http://mgi.nci.nih.gov/
1 (bases 1 to 910)
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
cDNA Library Preparation: Life Technologies, Inc.
cDNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM13606 row: e column: 19
High quality sequence stop: 577.
Location/Qualifiers
1..910
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6196986"
/sex="male"
/tissue_type="sciatic nerve"
/dev_stage="adult, 70 yr"
/lab_host="DH10B"
/clone_lib="Lupski_sciatic_nerve"
/Note="Vector: PCW7-SPORT6 (Life Technologies); Site_1:
NotI; Site_2: SalI; cDNA made by oligo-dT priming.
Directionally cloned using the following adaptors:
5'-TCGACCCAGCGCTCCG-3' and
5'-GACTAGTTCATGATCGAGCGCGCCCT(15)-3'. Size selected >
1 kb for average insert length 1.87 kb. This is a primary
library, non-amplified. Library constructed by Life
Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
College of Medicine) and is available through Life
Technologies."
```

```
Query Match 48.3%; Score 732; DB 13; Length 910;
Best Local Similarity 94.4%; Pred. No. 3.4e-151;
Matches 849; Conservative 0; Mismatches 31; Indels 19; Gaps 8;

QY 100 GGCCCCAGAGCTACTAGAAACGGTGTGACCCGCCACCGGCGCAGAGCTGCTCTGTG 159
DB 1 GGCCCCAGAGCTACTAGAAACGGTGTGACCCGCCACCGGCGCAGAGCTGCTCTGTG 60

QY 160 ATCCCTGGCAGTACTTCAAGACTGCTCCAGCCCGGCTAGCTCTCAGCCAGGAAGCTT 219
DB 61 ATCCCTGGCAGTACTTCAAGACTGCTCCAGCCCGGCTAGCTCTCAGCCAGGAAGCTT 120

QY 220 TCCTTACAGAGCGGCAGAGGAGCTATCTTGGAGCGCAGGCTGGGCTTATGCCAG 279
```

```
Db 121 TCCCTACAGGAGCGGCCAGCAGGAAGCTATCTGGAGGCGCAGGCTGGCCCTTATGCCAGG 180
Qy 280 GGGCTGCGCAGCAGCATCTCCCCCGGCGCTGGCGGAGGCCACCATGAGTCCACACAC 339
Db 181 GGGCTGCGCAGCAGCATCTCCCCCGGCGCTGGCGGAGGCCACCATGAGTCCACACAC 240
Qy 340 GTGGCCATCTCAGATGCGAGAGCTGCGTGAGCTGAACCACTGACAGTCAAGTGCAGAGTGAG 399
Db 241 GTGGCCATCTCAGATGCGAGAGCTGCGTGAGCTGAACCACTGACAGTCAAGTGCAGAGTGAG 300
Qy 400 ATTGCAAGGTGCTTCCAAAGAGAGTACTGAGCAGTATGGCTTTCCAGTCCG 459
Db 301 ATTGCAAGGTGCTTCCAAAGAGAGTACTGAGCAGTATGGCTTTCCAGTCCG 420
Qy 520 CCTCCCCGAGAGGTCCAGGCTGCCAGGAGGAGCAGCAGCAAGCAGCTGCTGCCCTG 579
Db 421 CCTCCCCGAGAGGTCCAGGCTGCCAGGAGGAGCAGCAGCAAGCAGCTGCTGCCCTG 480
Qy 580 GAGCGGTGTACAGAGATGCGCATCTGAAGAGTGGACCAAGTGAATGTGTCTCAA 639
Db 481 GAGCGGTGTACAGAGATGCGCATCTGAAGAGTGGACCAAGTGAATGTGTCTCAA 540
Qy 640 CTGATCGAGGTCTCGGATGACCCAGCTGAGCAGACCTCTATTTGGTGTGAGCTCTG 699
Db 541 CTGATCGAGGTCTCGGATGACCCAGCTGAGCAGACCTCTATTTGGTGTGAGCTCTG 600
Qy 700 AGAAGGCGGCGCTCATGAGTGCCTGTGACAAGCCCTTCTCGGAGGAGC-AAAGTCG 758
Db 601 AGAAGGCGGCGCTCATGAGTGCCTGTGACAAGCCCTTCTCGGAGGAGC-AAAGTCG 660
Qy 759 CTTTACTCGGAGAGTCTATCTGGGCTCGAGTACTGACCTGCCAGAGATCGTCCA 818
Db 661 CTTTACTCGGAGAGTCTATCTGGGCTCGAGTACTGACCTGCCAGAGATCGTCCA 720
Qy 819 CA-GGGACATCAAGCCATCAAGCTGCTCC--TGGGGGATGATGGGCGACGTTGAGATCG 874
Db 721 CAGGGGACATCAAGCCATCAAGCTGCTCC--TGGGGGATGATGGGCGACGTTGAGATCG 780
Qy 875 CCGAC-----TTTGGCGTACGACACCAAGTTT---GAGGGGAAACAGCGCTCAGC-TGTCC 924
Db 781 TCGCCCGACATTTGGCGTACGACCAAGTTTGAAGGGGAGCAGCAGCTCAGCTTGTCC 840
Qy 925 AGCAGCGGGGAA--CCCGACATTCATGG--CCCGCAGGCGCATTTCTGATTCGGCC 979
Db 841 AGCAACCGGGGAAACCCCGACATTCATGGGCCCCCGGAGGCGCATTTCTGAATTCGGCC 899
```

```
RESULT 11
BQ277875
LOCUS BQ277875 1066 bp mRNA linear EST 07-MAY-2002
DEFINITION AGENCOURT_7048118 NIH_MGC_109 Homo sapiens cDNA clone IMAGE:5804754
5', mRNA sequence.
ACCESSION BQ277875
VERSION BQ277875.1 GI:20488083
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1066)
NIH-MGC http://mgi.nci.nih.gov/
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL Contact: Robert Strausberg, Ph.D.
COMMENT Email: cgabs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
```

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.llnl.gov>
Plate: LCM2042 row: n column: 19
High quality sequence stop: 626.
Location/Qualifiers
1..1066
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5804754"
/tissue_type="teratocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_109"
/note="Organ: Ovary; Vector: pOTB7; Site 1: EcoRI; Site 2: XhoI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGACAGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

FEATURES

source

ORIGIN

Query Match 47.7%; Score 723; DB 13; Length 1066;
Best Local Similarity 97.3%; Pred. No. 3.6e-149;
Matches 779; Conservative 0; Mismatches 15; Indels 7; Gaps 4;
Qy 685 GTGTTTACCTCTCGAGAAAGGGCGCTCATGCAAGTGCCTGTGCAAGCCCTTCTCG 744
Db 36 GTGTTTACCTCTCGAGAAAGGGCGCTCATGCAAGTGCCTGTGCAAGCCCTTCTCG 95
Qy 745 GAGGAGCAGCTCCCTCTTACCTCGGGAGCTATCTCGGCTTCGAGTACTTGCATGC 804
Db 96 GAGGAGCAGCTCCCTCTTACCTCGGGAGCTATCTCGGCTTCGAGTACTTGCATGC 155
Qy 805 CAGAAGATCTGCCAGGGACATCAAGCCATCCAACTTCTCTCGGGGATGATGGGCAC 864
Db 156 CAGAAGATCTGCCAGGGACATCAAGCCATCCAACTTCTCTCGGGGATGATGGGCAC 215
Qy 865 GTGAAGATCTCCGACTTTGGCGTCAGCAACAGTTTTCAGGGGAAACAGCGCTCAGCTGCC 924
Db 216 GTGAAGATCTCCGACTTTGGCGTCAGCAACAGTTTTCAGGGGAAACAGCGCTCAGCTGCC 275
Qy 925 AGCAGCGGGGAAACCCAGCATTCATGCGCCCGGAGGCCATTTCTGATTCGGGCGAGC 984
Db 276 AGCAGCGGGGAAACCCAGCATTCATGCGCCCGGAGGCCATTTCTGATTCGGGCGAGC 335
Qy 985 TTCAAGTGGGAGGCGCTTGGATGATGGGCGCTACGTTGCTGCTGCTGCTGCTAT 1044
Db 336 TTCAAGTGGGAGGCGCTTGGATGATGGGCGCTACGTTGCTGCTGCTGCTGCTAT 395
Qy 1045 GGAAGTGGCCATTCATTCAGCATTTTCATCTCGCCCTCCACAGGAAGATCAAGATGAG 1104
Db 396 GGAAGTGGCCGTTTCATTCAGCATTTTCATCTCGCCCTCCACAGGAAGATCAAGATGAG 455
Qy 1105 CCGTGTGTGTTTCTGAGGAGCCAGAAATCAGCAGGAGCTCAAGACCTGATCTCTGAAG 1164
Db 456 CCGTGTGTGTTTCTGAGGAGCCAGAAATCAGCAGGAGCTCAAGACCTGATCTCTGAAG 515
Qy 1165 ATGTTAGCAAGAATCCGAGAGCAGAAATTTGGGTGCCAGACATCAAGTTGACCCCTGG 1224
Db 516 ATGTTAGCAAGAATCCGAGAGCAGAAATTTGGGTGCCAGACATCAAGTTGACCCCTGG 575
Qy 1225 GTGACCAAGAACGGGAGGAGCCCTTCTCTCGGAGGAGGACCTCCAGCGTGTGGAG 1284
Db 576 GTGACCAAGAACGGGAGGAGCCCTTCTCTCGGAGGAGGACCTCCAGCGTGTGGAG 635
Qy 1285 GTGACAGAGGGGAGGTTAAGACTCAGTCAGCTCATCCAGCTGGACACCGTGTATC 1344
Db 636 GTGACAGAGGGAGGTTAAGAACTCAGTCAGCTCATCCCGAGCTGAGACCGGATGC 695
Qy 1345 CTGGTGAAGTCCATGCTG-AGGAAGCGTTTCTTTGGGAACCCCTTTGAGCCCCAGGCAGC 1403

```
Db      696  CTGGGGAAGTCCATGCTGAAGGAAGCGTTCTTTGGGAACCCCTTTGAGCCCAAGCAG 755
|||
QY      1404 GAGGGAAGCCATCAGTGC--TGCTCCAGAAA--CCTACTGTGGAAGAGGG--TT 1457
|||
Db      756  GAGGGAAGGAGCATCATGCTCTGCTCCAGGAAACCTACTTGGTGAAGAAGGGGTTT 815
|||
QY      1458  TGGTGAAGGGGGCAAGAGCCC 1478
|||
Db      816  GGTGAAGGGGGCAAGAGCCC 836
|||

BX459685      972 bp      mRNA      linear      EST 22-MAY-2003
BX459685 Homo sapiens FETAL BRAIN Homo sapiens cDNA clone
CS0DF003YD04 5-PRIME, mRNA sequence.
BX459685
BX459685.1 GI:31035114
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 8786.r For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DF003DB02QP1&cluster=8786.r. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Paraday Avenue Genoscope sequence ID : CS0DF003DB02QP1.
Location/Qualifiers
1..972
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DF003YD04"
/tissue_type="FETAL BRAIN"
/dev_stage="fetal"
/clone_lib="Homo sapiens FETAL BRAIN"
/notes="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-strand cDNA was digested with Not I and
cloned into the Not I and EcoRV sites of the pCMVSPORT 6
vector. Library was not normalized."

ORIGIN
Query Match 47.7%; Score 722; DB 13; Length 972;
Best Local Similarity 97.2%; Pred. No. 5.7e-149;
Matches 754; Conservative 0; Mismatches 0; Indels 22; Gaps 1;

QY      1  ATGAGAGGGGGTCCAGTGTCTGTCGACGAGTCTCTGGGCAGAGCTGTGTAGAACGGGTG 60
|||
Db      197  ATGAGAGGGGGTCCAGTGTCTGTCGACGAGTCTCTGGGCAGAGCTGTGTAGAACGGGTG 256
|||
QY      61  GCAGCCATCATGTGACTCACTTGGAGAGGAGAGTGGTGGCCAGAGCCCTACTAGAAC 120
|||
Db      257  GCAGCCATCATGTGACTCACTTGGAGAGGAGAGTGGTGGCCAGAGCCCTACTAGAAC 316
|||
QY      121  GGTGTGACCCCCACACAGGGCCAGAGTGCCTCTGTGATCCTGCGAGTCTTCAAGA 180
|||
Db      317  GGTGTGACCCCCACACAGGGCCAGAGTGCCTCTGTGATCCTGCGAGTCTTCAAGA 376
|||
QY      181  CTGCTCCCGCGGCGCTAGCCTCTCAGCCAGGAAGCTTTCCCTACAGGAGCGGCCAGCA 240
|||
```

```
Db      377  CTGCTCCCGCGGCGCTAGCCTCTCAGCCAGGAAGCTTTCCCTACAGAGGGCGCAGCA 436
|||
QY      241  GGAAGCTATCTGGAGCGCAGGCTGGGCTTATGCCAGGGGCCCTGCCAGCCACATCTCC 300
|||
Db      437  GGAAGCTATCTGGAGCGCAGGCTGGGCTTATGCCAGGGGCCCTGCCAGCCACATCTCC 496
|||
QY      301  CCCCAGGCTCTGGCGGAGGCCACCATCGAGTCCCCACCAAGCTGGCCCATCTCAGATGCAGAG 360
|||
Db      497  CCCCAGGCTCTGGCGGAGGCCACCATCGAGTCCCCACCAAGCTGGCCCATCTCAGATGCAGAG 556
|||
QY      361  GACTCGTGCAGCTGAAACAGTACAGCTGCAGAGTGCAGATGGCAAGGTCCTACGGT 420
|||
Db      557  GACTCGTGCAGCTGAAACAGTACAGCTGCAGAGTGCAGATGGCAAGGTCCTACGGT 616
|||
QY      421  GTGGTGAGGCTGGCTTACAAACGAAAGTGAAGACAGACACTATGCAATGAAAGTCTCTTCC 480
|||
Db      617  GTGGTGAGGCTGGCTTACAAACGAAAGTGAAGACAGACACTATGCAATGAAAGTCTCTTCC 676
|||
QY      481  AAAAAAGAAAGTTACTGAAGCAGTATGGCTTTCACGTCGCCCTCCCCCGAGAGGGTCCCAG 540
|||
Db      677  AAAAAAGAAAGTTACTGAAGCAGTATGGCTTTCACGTCGCCCTCCCCCGAGAGGGTCCCAG 736
|||
QY      541  GTCGCCAGGAGGAGCCAGCCAGCAGCTGCTGCCCTGGAGGGGTGTACCAGGAGATT 600
|||
Db      737  GTCGCCAGGAGGAGCCAGCCAGCAGCTGCTGCCCTGGAGGGGTGTACCAGGAGATT 796
|||
QY      601  GCATCTCTGAAGAAGCTGGACCACTGAATGTGGTCAAACTGATCGAGTCTCTGGATGAC 660
|||
Db      797  GCATCTCTGAAGAAGCTGGACCACTGAATGTGGTCAAACTGATCGAGTCTCTGGATGAC 856
|||
QY      661  CCAGCTGAGGACCACTTATTTGGTTGTGACCTCTCTGAGAAAGG----- 706
|||
Db      857  CCAGCTGAGGACCACTTATTTGGTTGTGACCTCTCTGAGAAAGGCGGTGAGTTCCTCCCT 916
|||
QY      707  -----GGCCGCTCATGGAAGTGCCTGTGACAGCCCTCTCGAGAGCAAG 754
|||
Db      917  CCTGATCAGGCGCGTCATGGAAGTGCCTGTGACAGCCCTCTCTCGAGAGCAAG 972
|||

RESULT 13
BG396486      893 bp      mRNA      linear      EST 12-MAR-2001
602459214F1 NIH_MGC_16 Homo sapiens cDNA clone IMAGE:4581651 5',
mRNA sequence.
ACCESSION      BG396486
VERSION      BG396486.1 GI:13289845
KEYWORDS      EST.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NIH-MGC http://mgi.nhl.nih.gov/;
1 (bases 1 to 893)
AUTHORS      National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE      Unpublished (1999)
JOURNAL      Contact: Robert Strausberg, Ph.D.
COMMENT      Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: LLCM1302 row: 1 column: 04
High quality sequence stop: 733.
Location/Qualifiers
1..893
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4581651"
/tissue_type="retinoblastoma"
```


Db 314 ATTTTCATCTGGCCCTCCACAGGAAGATCAAGAAATGAGCCCGTGTGTTTCTGTAGGAGC 373
QY 1127 CAGAAATCAGCGAGAGCTCAAGGACCTGATCTGAAGATGTTAGACAAGAAATCCCGAGA 1186
Db 374 CAGAAATCAGCGAGAGCTCAAGGACCTGATCTGAAGATGTTAGACAAGAAATCCCGAGA 433
QY 1187 CGAAATTTGGGTGTCAGACATCAAGTTGACCTTTGGGTGACCAAGAACCGGGAGGAGC 1246
Db 434 CGAAATTTGGGTGTCAGACATCAAGTTGACCTTTGGGTGACCAAGAACCGGGAGGAGC 493
QY 1247 CCTTCCTTCGAGAGAGAGAGCTCCAGCGTGTGAGGTGACAGAGGGGAGGTTAAGA 1306
Db 494 CCTTCCTTCGAGAGAGAGAGACTCCAGCGTGTGAGGTGACAGAGGGAGGTTAAGA 553
QY 1307 ACTCAGTCAGGCTCATCCCAAGCTGAGACCAAGCTGATCTGTGAAGTCCATGCTGAGGA 1366
Db 554 ACTCAGTCAGGCTCATCCCAAGCTGAGACCAAGCTGATCTGTGAAGTCCATGCTGAGGA 613
QY 1367 AGCGTTCCTTTGGGAACCGTTTGGAGCCCGAGGACGAGGAGGAGGATCCATGTCG 1426
Db 614 AGCGTTCCTTTGGGAACCGTTTGGAGCCCGAGGACGAGGAGGAGGATCCATGTCG 673
QY 1427 CTCACGAAACCTACTGTGGAAGAGGTTTGGTGAAGGGGGGAGAGCCAGAGCTCC 1486
Db 674 CTCACGAAAC--TACTGTGGAAGAGGTTTGGTGAAGGGGGGAGAG--CCAAAGCTCC 730
QY 1487 CGCGGCTCCAGGAAGACGAGGCTGCATCC 1515
Db 731 CGCGGCTCCAGGAAGACAGGCTGATCC 759

RESULT 15
LOCUS BQ571613
DEFINITION UI-M-FCO-byc-e-19-0-UI.r1 NIH_BMAP_FCO Mus musculus cDNA clone
IMAGE:5716218 5', mRNA sequence.
ACCESSION BQ571613
VERSION BQ571613.1 GI:21474930
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 784)
AUTHORS NIH-MGC <http://imgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Dr. James Lin, University of Iowa
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LINL at:
<http://image.llnl.gov>
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)

Seq primer: pYX-5.
FEATURES
Location/Qualifiers
1..784
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:5716218"
/tissue_type="whole brain"
/dev_stage="embryo 12.5 dpc"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH_BMAP_FCO"
/notes="Organ: brain; Vector: pYX-Asc; Site 1: EcoR I;
Site 2: Not I; The library was constructed according to
Bonaldo, Lennon and Soares, Genome Research, 6:791-806,

1996. Denatured mRNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with Not I, and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail, is TCAGAGAGCC. This library was created for the University of Iowa Mouse Brain Molecular Anatomy Project (BMAP). 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institutes of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

ORIGIN

Query Match 44.1%; Score 668.6; DB 13; Length 784;
Best Local Similarity 91.6%; Pred. No. 3.4e-137;
Matches 718; Conservative 0; Mismatches 65; Indels 1; Gaps 1;
QY 261 GCGTGGGCTTATGTCACGGGGCTTCGAGCCACATCTCCCCCGGGGCTGCGGAGGCC 320
Db 2 GGTGGGCTTACTCTACAGGACCTGCCAGTCACATCTCTCTCGGTCTCTGCGGAGACC 61
QY 321 CACCATCGAGTCCCAACACCGTGGCCATCTCAGATCAGAGGACTCGTGCAGCTGAACCA 380
Db 62 CACCATCGAGTCCCAACCGTGGCCATCTCAGACAGAGGACTGTGTCAACTGAACCA 121
QY 381 GTACAAGCTTCAGAGTGAGATTGGCAAGGGTCCCTACGGTGTGGTGAAGCTGCCTTACAA 440
Db 122 GTACAAGCTTCAGAGTGAGATTGGCAAGGGTCCCTATGTTGGTGGTGGCTGTACAA 181
QY 441 CGAAAGTGAGACAGACACTATGCAATGAAAGTCTCTTCCAAAAGAAAGTTACTGAGCA 500
Db 182 CGAAAGTGAGACAGACACTATGCAATGAAAGTCTCTTCCAAAAGAAAGTTACTGAGCA 241
QY 501 GTATGGCTTTCCACGTGCGCCCTCCCGAGAGGGTCCAGGCTGCCAGGGAGGACGAC 560
Db 242 GTATGGCTTTCTCGCGCTCTCCCGAGAGGGTCCCAAGCTACCCAGGGAGGCGACG 301
QY 561 CAAGCAGCTGCTGCCCTCGAGAGGGGTGACCAGGAGATTGGCATCTCTGAAAGCTGGA 620
Db 302 CAACAGCTGCTGCCCTCGAGGGTGTGTATCAGGAGATTGCCATTCTAAAGCTGGA 361
QY 621 CCACGTGAATGTGTTCAAACTGATCGAGTCTCTGATGATGATGATGATGATGATGATGAT 680
Db 362 CCACGTGAATGTGTTCAAACTGATCGAGTCTCTGATGATGATGATGATGATGATGATGAT 421
QY 681 TTTGTGTTTGTACCTCTCGAGAAAGGGGCCCGTCTATGGAAGTGCCTGTGACAAGCCCTT 740
Db 422 TTTGTGTTTGTACCTCTCGAGAAAGGGGCCCGTCTATGGAAGTGCCTGTGACAAGCCCTT 481
QY 741 CTCGAGGAGCAAGCTCGCTCTTACCTCGGGAAGTCTATCTCTGGGCTCGAGTACTTGA 800
Db 482 CCCAGGAGCAAGCTCGCTCTTACCTCGGGAAGTCTATCTCTGGGCTCGAGTACTTGA 541
QY 801 CTGCCAAGAGATGTTGCCACAGGAGCAATCAAGCCATCAACCTGCTCTCTGGGAGATGAG 860
Db 542 CTGCCAAGAGATGTTGCCACAGGAGCAATCAAGCCATCAACCTGCTCTCTGGGAGATGAG 601
QY 861 GCACGTGAAGATCGCGACTTTGGGCTGAGCAACAGTTTGGAGGGAACAGCTCAGCT 920
Db 602 GCATGTGAAGATCGCGACTTTGGGCTGAGTAAACAGTTTGGAGGGAATGATGCTCAGCT 661
QY 921 GTCCAGCAGCGGGGAACCCCGAGCATTTATGGCCCCCGAGGCGCATTTCTGATTCGGGCA 980
Db 662 GTCCAGTACGGCAGGAGACCCCGCATTTATGGCCCCCGAGGCGCATTTCTGATTCGGGCA 721
QY 981 GAGCTTCAGTGGGAAGGCTTGGATGATGAGGCACTGGGTCACCTGTTACTGCTTTGT 1040
Db 722 GAGCTTCAGTGGGAAGG-CTTGGATGATGGGCACTGGGCTCACGCTATATGCTTTGT 780
1041 CTAT 1044
781 CTAT 784

Search completed: July 25, 2004, 11:34:37
Job time : 6003 secs

~~This Page Blank (uspio)~~

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 25, 2004, 06:22:55 ; Search time 140 Seconds
(without alignments)
6005.361 Million cell updates/sec

Title: US-10-690-617-1_COPY_173_1687

Perfect score: 1515

Sequence: 1 atggagggggtccagctgt.....aggaagacgaggtgcattcc 1515

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA.*

1: /cgn2_6/ptodata/2/ina/5A_COMB.seq.*

2: /cgn2_6/ptodata/2/ina/5B_COMB.seq.*

3: /cgn2_6/ptodata/2/ina/6A_COMB.seq.*

4: /cgn2_6/ptodata/2/ina/6B_COMB.seq.*

5: /cgn2_6/ptodata/2/ina/PTUS_COMB.seq.*

6: /cgn2_6/ptodata/2/ina/backfiles.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1515	100.0	2190	US-09-729-995-1	Sequence 1, Appli
2	1515	100.0	2190	US-10-135-689-1	Sequence 1, Appli
3	361	23.8	29629	US-09-729-995-3	Sequence 3, Appli
4	361	23.8	29629	US-10-135-689-3	Sequence 3, Appli
5	142.8	9.4	288	US-09-016-434-1008	Sequence 1008, Ap
6	99.4	6.6	236	US-09-016-434-809	Sequence 809, Appl
7	90.4	6.0	1302	US-08-913-050A-2	Sequence 2, Appli
8	90.4	6.0	1302	US-09-016-434-1146	Sequence 1146, Ap
9	87	5.7	1466	US-08-749-902-2	Sequence 2, Appli
10	83.8	5.5	2968	US-08-685-852-1	Sequence 1, Appli
11	83.4	5.5	1599	US-09-256-465-1	Sequence 1, Appli
12	83.4	5.5	1599	US-09-167-322-3	Sequence 3, Appli
13	83.4	5.5	1599	US-09-023-655-1004	Sequence 1004, Ap
14	82.8	5.5	1257	US-09-799-875-15	Sequence 15, Appli
15	82.8	5.5	1826	US-09-799-875-13	Sequence 13, Appli
16	82.8	5.5	1864	US-09-819-607-1	Sequence 1, Appli
17	82.8	5.5	1975	US-08-852-743-1	Sequence 1, Appli
18	82.8	5.5	1975	US-09-185-370-1	Sequence 1, Appli
19	82.2	5.4	1636	US-09-016-434-1433	Sequence 1433, Ap
20	79.4	5.2	2158	US-09-765-815-1	Sequence 1, Appli
21	79.4	5.2	3604	US-09-688-188B-27	Sequence 27, Appli
22	79.4	5.2	3604	US-09-291-417D-27	Sequence 27, Appli
23	78.6	5.2	1940	US-09-718-032-1	Sequence 1, Appli
24	78.6	5.2	1941	US-09-082-737-1	Sequence 1, Appli
25	78.6	5.2	2050	US-09-688-188B-28	Sequence 28, Appli
26	78.6	5.2	2050	US-09-291-417D-28	Sequence 28, Appli
27	78.6	5.2	2806	US-09-688-188B-102	Sequence 102, Appl

ALIGNMENTS

RESULT 1

US-09-729-995-1
; Sequence 1, Application US/09729995
; Patent No. 6426206
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL000904
; CURRENT APPLICATION NUMBER: US/09/729,995
; CURRENT FILING DATE: 2000-12-06
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2190
; TYPE: DNA
; ORGANISM: Human
US-09-729-995-1

Query Match	100.0%;	Score 1515;	DB 4;	Length 2190;	
Best Local Similarity	100.0%;	Pred. No. 0;			
Matches 1515;	Conservative	0;	Mismatches	0;	
		Indels	0;	Gaps	0;
QY	1	ATGAGGGGGGTCCAGCTGTCTGCTGCAGGATCCTCGGGCAGAGCTGGTAGAACGGGTG	60		
Db	173	ATGAGGGGGGTCCAGCTGTCTGCTGCAGGATCCTCGGGCAGAGCTGGTAGAACGGGTG	232		
QY	61	CGAGCCATCGATGTGACTCACTTGGAGGAGCGAGATGGTGGCCAGAGCCTACTAGAAAC	120		
Db	233	CGAGCCATCGATGTGACTCACTTGGAGGAGCGAGATGGTGGCCAGAGCCTACTAGAAAC	292		
QY	121	GGTGTGAGACCCCAACACCGGGCGAGAGCTGCCTCTGTGATCCCTGGCAGTACTTCAAGA	180		
Db	293	GGTGTGAGACCCCAACACCGGGCGAGAGCTGCCTCTGTGATCCCTGGCAGTACTTCAAGA	352		
QY	181	CTGCTCCCAAGCGGCGCTAGCCTCTCAGCCAGGAGAGCTTTCCTTACAGGAGCGGCCAGCA	240		
Db	353	CTGCTCCCAAGCGGCGCTAGCCTCTCAGCCAGGAGAGCTTTCCTTACAGGAGCGGCCAGCA	412		
QY	241	GGAAGCTATCTGGAGGCGCAGGCTGGGCCCTTATGCCACGGGGCTGCCAGCCACATCTCC	300		
Db	413	GGAAGCTATCTGGAGGCGCAGGCTGGGCCCTTATGCCACGGGGCTGCCAGCCACATCTCC	472		
QY	301	CCCCGGGCGCTGGGGAGGCCCAACCATCGAGTCCCAACACGCTGGCCATCTCAGATGCAGAG	360		
Db	473	CCCCGGGCGCTGGGGAGGCCCAACCATCGAGTCCCAACACGCTGGCCATCTCAGATGCAGAG	532		
QY	361	GACTGCGTGCAGCTGAACCAAGTCAAGCTCGAGTGAGATTGGCAAGGGTGCCCTACGGT	420		

Db 533 GACTGCCTGCAGCTGAACCACTACAAGCTGCAGAGTGAGATTGCAAGGGTGCCCTACGGT 592
Qy 421 GTGTGTGGCTGGCTCTCAACGAAGTGAAGACAGACATATGCAATGAAGTCCCTTTCC 480
Db 593 GTGTGTGGCTGGCTCTCAACGAAGTGAAGACAGACATATGCAATGAAGTCCCTTTCC 652
Qy 481 AAAAAGAGTTACTGAAGCAGTATGGCTTTCCAGTGGCCCTCCCGAGAGGGTCCAG 540
Db 653 AAAAAGAGTTACTGAAGCAGTATGGCTTTCCAGTGGCCCTCCCGAGAGGGTCCAG 712
Qy 541 GCTGCCAGGAGACCAAGCAGCTGCTGCCCTCTGAGCGGGTGTCACGAGATT 600
Db 713 GCTGCCAGGAGACCAAGCAGCTGCTGCCCTCTGAGCGGGTGTCACGAGATT 772
Qy 601 GCCATCTGAAGAGCTGGACCAAGTGAATGTGTCAAACTGATCGAGGCTCTGGATGAC 660
Db 773 GCCATCTGAAGAGCTGGACCAAGTGAATGTGTCAAACTGATCGAGGCTCTGGATGAC 832
Qy 661 CCAGCTGAGCAACCTCTATTGGTGTGTTGACCTCTGAGAAAGGGGCCGTCATGAA 720
Db 833 CCAGCTGAGCAACCTCTATTGGTGTGTTGACCTCTGAGAAAGGGGCCGTCATGAA 892
Qy 721 GTGCCCTGTGCAAGCCCTCTTCGGAGGAGCAAGCTCGCCCTCTACCTCGCGGAGCTCATC 780
Db 893 GTGCCCTGTGCAAGCCCTCTTCGGAGGAGCAAGCTCGCCCTCTACCTCGCGGAGCTCATC 952
Qy 781 CTGGGCTCGAGTACTTGCTGCTGCAAGATCGTCCACAGGACATCAAGCCATCCAAC 840
Db 953 CTGGGCTCGAGTACTTGCTGCTGCAAGATCGTCCACAGGACATCAAGCCATCCAAC 1012
Qy 841 CTGCTCTCGGGGATGATGGCAGCTGAAGATCGCGACTTGTGCGTCAGCAACAGTTT 900
Db 1013 CTGCTCTCGGGGATGATGGCAGCTGAAGATCGCGACTTGTGCGTCAGCAACAGTTT 1072
Qy 901 GAGGGGAACAGCTGCTAGCTGTCCAGCACGGCGGGAAACCCAGCATTCATGGCCCCCGAG 960
Db 1073 GAGGGGAACAGCTGCTAGCTGTCCAGCACGGCGGGAAACCCAGCATTCATGGCCCCCGAG 1132
Qy 961 GCCATTCTGATTCGGGCGAGAGCTTCAGTGGGAGSCCTTGATGTATGGGCGACTGGC 1020
Db 1133 GCCATTCTGATTCGGGCGAGAGCTTCAGTGGGAGSCCTTGATGTATGGGCGACTGGC 1192
Qy 1021 GTCAAGTTGACTGCTTTGCTATGGGAAGTGCCATTCTATGACAGATTTTCATCTCGCC 1080
Db 1193 GTCAAGTTGACTGCTTTGCTATGGGAAGTGCCATTCTATGACAGATTTTCATCTCGCC 1252
Qy 1081 CTCACAGGAAGATCAAGATAGCCGCTGTTTCTCAGGAGCCAGAAATCAGCGAG 1140
Db 1253 CTCACAGGAAGATCAAGATAGCCGCTGTTTCTCAGGAGCCAGAAATCAGCGAG 1312
Qy 1141 GAGCTCAAGCACTGATCTCGAGATGTTAGACAGATATCCGAGAGGAGTGGGTTG 1200
Db 1313 GAGCTCAAGCACTGATCTCGAGATGTTAGACAGATATCCGAGAGGAGTGGGTTG 1372
Qy 1201 CCAGACATCAAGTTGCAACCTTCGGGTGACCAAGAACGGGAGGAGCCCTTCCTTCGGAG 1260
Db 1373 CCAGACATCAAGTTGCAACCTTCGGGTGACCAAGAACGGGAGGAGCCCTTCCTTCGGAG 1432
Qy 1261 GAGGAGCACTGACGCTGTGGAGTGACAGAGGGGAGGTTAAGAACTCAGTCAGGCTC 1320
Db 1433 GAGGAGCACTGACGCTGTGGAGTGACAGAGGGGAGGTTAAGAACTCAGTCAGGCTC 1492
Qy 1321 ATCCCGAGCTGGACCAAGCTGATCTGCTGAAGTCCATGCTGAGGAGCGTCTCTTTGG 1380
Db 1493 ATCCCGAGCTGGACCAAGCTGATCTGCTGAAGTCCATGCTGAGGAGCGTCTCTTTGG 1552
Qy 1381 AACCCGTTTGAAGCCCGAGGCAACGAGGAGAGCGATCCATGCTCTCCAGGAACCTA 1440
Db 1553 AACCCGTTTGAAGCCCGAGGCAACGAGGAGAGCGATCCATGCTCTCCAGGAACCTA 1612
Qy 1441 CTGGTGAAGAGGGTTTGGTGAAGGGGCAAGGCGCAGAGCTCCCGGCTCCAGGA 1500
Db 1613 CTGGTGAAGAGGGTTTGGTGAAGGGGCAAGGCGCAGAGCTCCCGGCTCCAGGA 1672

Qy 1501 GACGAGCTGCATCC 1515
Db 1673 GACGAGCTGCATCC 1687

RESULT 2

US-10-135-689-1
; Sequence 1, Application US/10135689
; Patent No. 6670162
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL000904DIV
; CURRENT APPLICATION NUMBER: US/10/135,689
; CURRENT FILING DATE: 2002-05-01
; PRIOR APPLICATION NUMBER: 60/247,031
; PRIOR FILING DATE: 2000-11-13
; PRIOR APPLICATION NUMBER: 09/729,995
; PRIOR FILING DATE: 2000-12-06
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2190
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-135-689-1

Query Match 100.0%; Score 1515; DB 4; Length 2190;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1515; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGAGAGGGGGTCCAGCTGCTGCCAGGATCTTCGGCAGAGCTGTGAGAACGGTG 60
Db 173 ATGAGAGGGGGTCCAGCTGCTGCCAGGATCTTCGGCAGAGCTGTGAGAACGGTG 232
Qy 61 GCAGCCATCATGTGATCTCACTTGGAGGAGGAGATGGTGGCCAGAGCCCTACTAGAAC 120
Db 233 GCAGCCATCATGTGATCTCACTTGGAGGAGGAGATGGTGGCCAGAGCCCTACTAGAAC 292
Qy 121 GGTGTGACCCCCACACACGGCCAGAGCTGCTGTGATCCTGCGACTACTTCAAGA 180
Db 293 GGTGTGACCCCCACACGGCCAGAGCTGCTGTGATCCTGCGACTACTTCAAGA 352
Qy 181 CTGCTCCAGCCCGGCTAGCCCTCTCAGCCAGGAAGCTTTCCTACAGGAGCGGCCAGCA 240
Db 353 CTGCTCCAGCCCGGCTAGCCCTCTCAGCCAGGAAGCTTTCCTACAGGAGCGGCCAGCA 412
Qy 241 GGAAGCTATCTGGAGGCGGAGGCTGGGCTTATGCCACGGGGCTTCGACGCCATCTCC 300
Db 413 GGAAGCTATCTGGAGGCGGAGGCTGGGCTTATGCCACGGGGCTTCGACGCCATCTCC 472
Qy 301 CCCCAGGCTGGCGGAGGCCCACCATCTGAGTCCCAACACGTCGCCATCTCAGATGCAAG 360
Db 473 CCCCAGGCTGGCGGAGGCCCACCATCTGAGTCCCAACACGTCGCCATCTCAGATGCAAG 532
Qy 361 GACTGCGTCAGCTGAACCAAGCTCAAGCTGCAAGTGAGATTGGCAAGGGTGCTACGGT 420
Db 533 GACTGCGTCAGCTGAACCAAGCTCAAGCTGCAAGTGAGATTGGCAAGGGTGCTACGGT 592
Qy 421 GTGCTCAGGCTGGCTTACCAACGAAAGTGAAGACAGACATATGCAATGAAGTCCCTTCC 480
Db 593 GTGCTCAGGCTGGCTTACCAACGAAAGTGAAGACAGACATATGCAATGAAGTCCCTTCC 652
Qy 481 AAAAAGAGTTACTGAAGCAGTATGGCTTTCCACGTCGCGCTCCCGGAGAGGGTCCAG 540
Db 653 AAAAAGAGTTACTGAAGCAGTATGGCTTTCCACGTCGCGCTCCCGGAGAGGGTCCAG 712
Qy 541 GCTGCCAGGAGGACCAAGCAGAGCTGTGCGCCCTCGAGCGGGTGATCCAGAGATT 600
Db 713 GCTGCCAGGAGGACCAAGCAGAGCTGTGCGCCCTCGAGCGGGTGATCCAGAGATT 772

```
QY 601 GCCATCCTGAAGAGCTGCACACGCTGAATGTGGTCAAACTGATCGAGTCTCTGGATGAC 660
    |||||
Db 773 GCCATCCTGAAGAGCTGCACACGCTGAATGTGGTCAAACTGATCGAGTCTCTGGATGAC 832
QY 661 CCAGCTGAGGACAACTTATTTGGTGTGTTGACCTCTTGAGAAAGGGCCCGTCAITGAA 720
    |||||
Db 833 CCAGCTGAGGACAACTTATTTGGTGTGTTGACCTCTTGAGAAAGGGCCCGTCAITGAA 892
QY 721 GTGCCCTGTGACAAAGCCCTTCTCGAGAGCAAGCTCGCTTACCTCGGGAGCTCATC 780
    |||||
Db 893 GTGCCCTGTGACAAAGCCCTTCTCGAGAGCAAGCTCGCTTACCTCGGGAGCTCATC 952
QY 781 CTGGGCCCTCGAGTACTTTCACCTGCGAGAGATCGTCCACAGGACATCAAGCCATCCAAAC 840
    |||||
Db 953 CTGGGCCCTCGAGTACTTTCACCTGCGAGAGATCGTCCACAGGACATCAAGCCATCCAAAC 1012
QY 841 CTGCTCTCTGGGGGATGATGGGCACGTGAAGATCGCCGACTTTGGCGTCAAGCAACAGTTT 900
    |||||
Db 1013 CTGCTCTCTGGGGGATGATGGGCACGTGAAGATCGCCGACTTTGGCGTCAAGCAACAGTTT 1072
QY 901 GAGGGGACGACGCTCAGCTGTCCAGCACGGCGGGAACCCAGCATTCATGGCCCCCGAG 960
    |||||
Db 1073 GAGGGGACGACGCTCAGCTGTCCAGCACGGCGGGAACCCAGCATTCATGGCCCCCGAG 1132
QY 961 GCATTTCTGATTCGGGCCAGAGCTTTCAGTGGGAAGGCGCTTGGATGTATGGGCCACTGGC 1020
    |||||
Db 1133 GCATTTCTGATTCGGGCCAGAGCTTTCAGTGGGAAGGCGCTTGGATGTATGGGCCACTGGC 1192
QY 1021 GTCACTGTTGATCTGTTGTCTATGGGAAGTGCCTCATTCAGACGATTTCTCTGGCC 1080
    |||||
Db 1193 GTCACTGTTGATCTGTTGTCTATGGGAAGTGCCTCATTCAGACGATTTCTCTGGCC 1252
QY 1081 CTCACAGGAGATCAAGATGAGCCCGTGTGTTTCTTGAGAGCCAGAAATCAGCGAG 1140
    |||||
Db 1253 CTCACAGGAGATCAAGATGAGCCCGTGTGTTTCTTGAGAGCCAGAAATCAGCGAG 1312
QY 1141 GAGCTCAAGGACCTGATCTGAAGATGTTAGACAAGAAATCCCGACAGAGAAATGGGGTG 1200
    |||||
Db 1313 GAGCTCAAGGACCTGATCTGAAGATGTTAGACAAGAAATCCCGACAGAGAAATGGGGTG 1372
QY 1201 CCAGACATCAAGTTGACCCCTTGGGTGACCAAGAACGGGGAGGAGCCCTTCTCTCGGAG 1260
    |||||
Db 1373 CCAGACATCAAGTTGACCCCTTGGGTGACCAAGAACGGGGAGGAGCCCTTCTCTCGGAG 1432
QY 1261 GAGGAGCACTGACGCTGTGTGAGGTGACAGAGGGGGAGTTAAGAACTCAGTCAGGCTC 1320
    |||||
Db 1433 GAGGAGCACTGACGCTGTGTGAGGTGACAGAGGGGGAGTTAAGAACTCAGTCAGGCTC 1492
QY 1321 ATCCCCAGCTGACCAACGCTGATCTCTGTGAAGTCCATGCTGAGGAAGCGTTTCTTTGGG 1380
    |||||
Db 1493 ATCCCCAGCTGACCAACGCTGATCTCTGTGAAGTCCATGCTGAGGAAGCGTTTCTTTGGG 1552
QY 1381 AACCGCTTGTAGCCCCAGGACGAGGGAAGCGATCCATCTCTGCTCCAGGAACCTTA 1440
    |||||
Db 1553 AACCGCTTGTAGCCCCAGGACGAGGGAAGCGATCCATCTCTGCTCCAGGAACCTTA 1612
QY 1441 CTGTGTGAAGAGGGTTTGGTGAAGGGGCAAGAGCCAGAGCTCCCGGCGTCCAGGAA 1500
    |||||
Db 1613 CTGTGTGAAGAGGGTTTGGTGAAGGGGCAAGAGCCAGAGCTCCCGGCGTCCAGGAA 1672
QY 1501 GACGAGGCTGCATCC 1515
    |||||
Db 1673 GACGAGGCTGCATCC 1687
    |||||
```

RESULT 3

```
US-09-729-995-3
; Sequence 3, Application US/09729995
; Patent No. 6426206
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
```

```
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL000904
; CURRENT APPLICATION NUMBER: US/09/729,995
; CURRENT FILING DATE: 2000-12-06
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 29629
; TYPE: DNA
; ORGANISM: Human
; US-09-729-995-3
```

```
Query Match 23.8%; Score 361; DB 4; Length 29629;
Best Local Similarity 100.0%; Pred. No. 8e-83;
Matches 361; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAGGGGGGTCCAGCTGTCTGCTGCCAGGATCTCTCGGCAGAGCTGTAGAAAGGGTG 60
    |||||
Db 3000 ATGAGGGGGGTCCAGCTGTCTGCTGCCAGGATCTCTCGGCAGAGCTGTAGAAAGGGTG 3059
QY 61 CGAGCCATCGATGTGACTCACTTGGAGGAGCAGATGTGGCCAGAGCCTACTAGAAAC 120
    |||||
Db 3060 CGAGCCATCGATGTGACTCACTTGGAGGAGCAGATGTGGCCAGAGCCTACTAGAAAC 3119
QY 121 GGTGTGACCCCCACACACGGGCCAGAGCTGCCTCTGTGATCCCTGGCAGTACTTCAAGA 180
    |||||
Db 3120 GGTGTGACCCCCACACACGGGCCAGAGCTGCCTCTGTGATCCCTGGCAGTACTTCAAGA 3179
QY 181 CTGCTCCAGCCCGGCTTAGCTCTCAGCCAGGAAGCTTTCCTACAGGAGGGCCAGCA 240
    |||||
Db 3180 CTGCTCCAGCCCGGCTTAGCTCTCAGCCAGGAAGCTTTCCTACAGGAGGGCCAGCA 3239
QY 241 GSAAGCTATCTGGAGCGCAGCTGGGCTTATGCCAGGGGCTTCAGCCACATCTCC 300
    |||||
Db 3240 GSAAGCTATCTGGAGCGCAGCTGGGCTTATGCCAGGGGCTTCAGCCACATCTCC 3299
QY 301 CCCCCGGGCTCGCGGAGGGCCCCACCATCGAGTCCCAACACGTGGCCATCTCAGATGCAGAG 360
    |||||
Db 3300 CCCCCGGGCTCGCGGAGGGCCCCACCATCGAGTCCCAACACGTGGCCATCTCAGATGCAGAG 3359
QY 361 G 361
    |
Db 3360 G 3360
```

RESULT 4

```
US-10-135-689-3
; Sequence 3, Application US/10135689
; Patent No. 6670162
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL000904DIV
; CURRENT APPLICATION NUMBER: US/10/135,689
; CURRENT FILING DATE: 2002-05-01
; PRIOR APPLICATION NUMBER: 60/247,031
; PRIOR FILING DATE: 2000-11-13
; PRIOR APPLICATION NUMBER: 09/729,995
; PRIOR FILING DATE: 2000-12-06
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 29629
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-135-689-3
```

```
Query Match 23.8%; Score 361; DB 4; Length 29629;
Best Local Similarity 100.0%; Pred. No. 8e-83;
Matches 361; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 ATGAGGGGGTCCAGCTGTCTCTGCCAGGATCCTCGGCGAGAGCTGGTAGAACGGGTG 60
Db |||||||
3000 ATGAGGGGGTCCAGCTGTCTGTGCGAGGATCCTCGGCGAGAGCTGGTAGAACGGGTG 3059
QY 61 GCAGCCATCGATGTGACTCACTTGGAGGAGCAGATGTGTGCGGAGAGCTTACTAGAAC 120
Db |||||||
3060 GCAGCCATCGATGTGACTCACTTGGAGGAGCAGATGTGTGCGGAGAGCTTACTAGAAC 3119
QY 121 GGTGTGAGCCCCCACCACCGAGGCTGCTGTGTGATPCCCTGGCAGTACTTCAAGA 180
Db |||||||
3120 GGTGTGAGCCCCCACCACCGAGGCTGCTGTGTGATPCCCTGGCAGTACTTCAAGA 3179
QY 181 CTGCTCCAGCGCGCTAGCTCTCAGCCAGGAGCTTTCCTACAGGAGCGCCAGCA 240
Db |||||||
3180 CTGCTCCAGCGCGCTAGCTCTCAGCCAGGAGCTTTCCTACAGGAGCGCCAGCA 3239
QY 241 GGAAGCTATCTGGAGGCGCAGGCTGGGCTTATGCCAGCGGCGCTGCCAGCCATCTCC 300
Db |||||||
3240 GGAAGCTATCTGGAGGCGCAGGCTGGGCTTATGCCAGCGGCGCTGCCAGCCATCTCC 3299
QY 301 CCCCAGGCTGGGAGGCGCCACCATCGAGTCCACACAGTGGCCATCTCAGATGCAGAG 360
Db |||||||
3300 CCCCAGGCTGGGAGGCGCCACCATCGAGTCCACACAGTGGCCATCTCAGATGCAGAG 3359
QY 361 G 361
Db |
3360 G 3360
```

RESULT 5

```
US-09-016-434-1008
; Sequence 1008, Application US/09016434
; Patent No. 6500938
; GENERAL INFORMATION:
; APPLICANT: Janice Au-Young
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
; NUMBER OF SEQUENCES: 1490
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION NUMBER: US/09/016,434
; FILING DATE: HEREWITH
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0002 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 1008:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 288 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
```

```
; IMMEDIATE SOURCE:
; LIBRARY: BRAINOT04
; CLONE: 926034
; US-09-016-434-1008
Query Match 9.4%; Score 142.8; DB 4; Length 288;
Best Local Similarity 70.7%; Pred. No. 1.7e-27;
Matches 203; Conservative 0; Mismatches 83; Indels 1; Gaps 1;
QY 472 GTCTTTTCCAAAAGAAAGTTACTGAAGCAGTATGGCTTTTCCACGTGCGCTTCCCGCAGA 531
Db |||||||
1 GTGCTGTCCAAAAGAAAGTCTGATCCGCGAGCGCGCTTTCACGTGCGCTTCCACCCCGA 60
QY 532 GGGTCCAGGCTGCCAGGAGGAGCAGCCAGCAGCTGTGCGCTTGGAGCGGGGTGTAC 591
Db |||||||
61 GGCACCGCGGCGAGCTCCTGGAGGCTGATCAGCCAGGGGCGCCCATTTGAGCAGGTGTAC 120
QY 592 CAGGAGATTGCATCTCTGAAGAGCTGCACCATGTGAATGTGGTCAAACTGATCGAGTGC 651
Db |||||||
121 CAGGAATTGCGATCTCTCAAGAGCTGACCCCATTTGGTGAAGCTGGTGGAGGTC 180
QY 652 CTGATGACCCAGCTGAGGACCACTCTATTGGTGTGTTGACCTCTTGAAGAAAGGGGCC 711
Db |||||||
181 CTGATGACCCCAATGAGGACCATCTGTACATGGTGTTCGA-ATGGTCAACCAAGGGGCC 239
QY 712 GTCATGGAAGTGGCTCTGTGACAGCCCTTCTCGAGGAGCAAGCTCG 758
Db |||||||
240 GTGATGGAAGTGGCCACCCTTAACCCANTTTTGAAGACCAAGGGGCC 286
```

RESULT 6

```
US-09-016-434-809
; Sequence 809, Application US/09016434
; Patent No. 6500938
; GENERAL INFORMATION:
; APPLICANT: Janice Au-Young
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
; NUMBER OF SEQUENCES: 1490
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION NUMBER: US/09/016,434
; FILING DATE: HEREWITH
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0002 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 809:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 236 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
```



```
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/016,434
; FILING DATE: HEREWITH
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0002 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 1146:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1302 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GENBANK
; CLONE: 91480860
; US-09-016-434-1146

Query Match          6.0%; Score 90.4; DB 4; Length 1302;
Best Local Similarity 50.6%; Pred. No. 9.7e-14;
Matches 335; Conservative 0; Mismatches 306; Indels 21; Gaps 4;

Qy 586 GTGTACCAGGAGATTGCCATCCTGAAGAAGCTGGACCACTGTAATGTGGTCAAACTGATC 645
Db |||
Qy 283 GTGAAGAAGGAAATTCAACTACTGAGGAGTTACGGCAAAAATGTATCCAGCTGGTG 342
Db |||
Qy 646 GAGCTCTGTGATGACCCAGCTGAGGACAACTCTATTGGTGTGTA-----CCTCTG 699
Db |||
Qy 343 GATGTGTATACACGAGAGAGAGAGAAATGTATGTGTGATGGAGTACTGGTGTGT 402
Db |||
Qy 700 AGAAGGGGGCCGCTCATGGAAGTGCCTGTGACAAAGCCCTTCTCGGAGGACAGCTCGC 759
Db |||
Qy 403 GGCATGACGAGAAATGTGGACAGGTGCGGAGAGCGTTTCCAGGTGTGCGAGCCAC 462
Db |||
Qy 760 CTCTACTCGGGAGCTCATCCTGGGCTCGAGTACTTGCTACCTCCAGAGATCGTCCAC 819
Db |||
Qy 463 GGGTACTTCTGTACGTGATGACGGCTGGAGTACCTGATAGCCAGGGCAATTTGCAC 522
Db |||
Qy 820 AGGACATCAAGCCATCCAACTGCTCTGGGGATGATGGGACAGTGAATGCCGAC 879
Db |||
Qy 523 AAGGACATCAAGCGGGAACTCTGTCTACCAACCGGTGGCACCCCTCAAAATCTCCGAC 582
Db |||
Qy 880 TTGGCGCTCAGCAACCACTTTGAGGGGAACAGCCTCAGCTGTCCAGCAGGC-----G 933
Db |||
Qy 583 CTGGCGGTGGCGAGGCACTGCACCCGTTGCGGCGGAGACACCTGCGGACAGCCAG 642
Db |||
Qy 934 GGAACCCAGCATTCATGGCCCGGAGGCCATTTCTGATTCGGCCAGAGCTTCAGTGGG 993
Db |||
Qy 643 GGCTCCCGCGCTTTCAGCGCCCGCGAG---ATTGCCAAAGCCCTGGGACACCTTCTCGGC 699
Db |||
Qy 994 AAGGCTTGTGATGATGGGCACTGGCGTCAAGTGTACTGCTTTGTCTATGGAAGTGC 1053
Db |||
Qy 700 TTCAAGTGAACAATCTGCTGGGTGGGTCAACCTCTACAAACATCACACGGGTCTGTAC 759
Db |||
Qy 1054 CCATTTCATCGAGATTTCATCTTGGCCCTCCACAGGAAGATCAAGATGAGCCCGTGGTG 1113
Db |||
Qy 760 CCCTTCGAAGGGGACAACTCTACAAGTTGTTGAGAACATCGGAAGGGAGCTAGCC 819
Db |||
Qy 1114 TTCTCTGAGAGCCAGAAATCAGGAGGAGCTCAAGGACTGATPCTTGAAATGTTAGAC 1173
Db |||
Qy 820 ATCCCGGGGAGTGTG-----GCCCGCCGCTCTCTGACCTGTGAAAGGGATGCTTGAG 873
Db |||
Qy 1174 AAGAAATCCCGAGACGAGAAATTTGGGGTGCAGACATCAAGTTGCACCCCTGGGTGACCAAG 1233
Db |||
Qy 874 TAGCAACCGGCCAAGAGGTTCTCCATCCGGCAGATCGGAGCAGCAGCTGGTTCGGGAAG 933
Db |||
```

```
Qy 1234 AA 1235
Db ||
Db 934 AA 935

RESULT 9
US-08-749-902-2
; Sequence 2, Application US/08749902
; Patent No. 5985635
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Goli, Surya K.
; APPLICANT: Hillman, Jennifer L.
; TITLE OF INVENTION: NOVEL HUMAN SERINE/THREONINE
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: US
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/749,902
; FILING DATE: Filed Herewith
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0150 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1466 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: Consensus
; US-08-749-902-2
```

```
Query Match          5.7%; Score 87; DB 2; Length 1466;
Best Local Similarity 50.0%; Pred. No. 7.7e-13;
Matches 331; Conservative 0; Mismatches 310; Indels 21; Gaps 4;
```

```
Qy 586 GTGTACCAGAGATTGCCATCCTGAAGAAGCTGGACCACTGATGTGGTCAAACTGATC 645
Db |||
Qy 382 GTGAAGAAGAAATTCAACTACTGAGGAGTTACGGCAGAAAATGTATCCAGCTGGTG 441
Db |||
Qy 646 GAGGTCTCTGGATGACCCAGCTGAGGACAACTCTATTGGTGTGTA-----CCTCTG 699
Db |||
Qy 442 GATGTGTATACACGAGAGAGAGAGAAATGTATGTGATGAGTACTGCGTGTGT 501
Db |||
Qy 700 AGAAGGGGCGGCTCATGGAAGTGCCTGTGAACGCCCTTCTCGGAGGAGCAAGCTGCG 759
Db |||
Qy 502 GGCATGAGGAATGTGGACACAGCTGCGGAGAGCGTTTCCAGTGTGCCAGGCCAC 561
Db |||
Qy 760 CTCTACTCGGGACCTCATCTCTGGCCCTGAGTACTTGCACTGCCAGAGATCGTCCAC 819
Db |||
Qy 562 GGGTACTTCTGTGCTGATGATGACGNTGGGGTACTCTGATNGCCAGGNAATTTGTGAC 621
Db |||
Qy 820 AGGACATCAAGCCATCCCAACCTGTCTCTGGGGGATGATGGGCACTGGAAGATCGCCGAC 879
Db |||
```


Db 1198 CGCAGCTCAGCCCGAGCCAAAGTCCCTGCTGCTGAGGCTGCTTAAGAGGACCCCAAG 1257
QY 1186 ACGAATTTGGGTG 1200
Db 1258 CAGAGGCTTGGTGG 1272

RESULT 12

US-09-167-322-3
; Sequence 3, Application US/09167322
; Patent No. 6365151
; GENERAL INFORMATION:
; APPLICANT: Allegheny University of the Health
; Sciences, Halpern, Michael S.
; England, James M.
; TITLE OF INVENTION: CANCER VACCINE
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seidel, Gonda, Lavorigna & Monaco, P.C.
; STREET: Suite 1800, Two Penn Center Plaza
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19102

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/167,322
; FILING DATE: 07-Oct-1998
; CLASSIFICATION: <unknown>

PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US97/00582
; FILING DATE: <unknown>

ATTORNEY/AGENT INFORMATION:
; NAME: Monaco, Daniel A.

REGISTRATION NUMBER: 30,480
; REFERENCE/DOCKET NUMBER: 7933-33 PC

TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-8383

TELEFAX: (215) 568-5549

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 1599 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 3:

US-09-167-322-3

Query Match 5.5%; Score 83.4; DB 4; Length 1599;
Best Local Similarity 51.3%; Pred. No. 6.7e-12;
Matches 254; Conservative 0; Mismatches 226; Indels 15; Gaps 2;

QY 706 GGGCCCGCTCATGGAGTGCGCTGTGACAGCCCTTCTCGGAGGAGCAAGCTCGCCTCTAC 765

Db 793 GAGCTGTCTTCCACCTGTCCGGGAGCGGTCTTTCACAGAGGAGCGGCCCGGTTTAT 852

QY 766 CTGGGGGACGTATCTCGGGCGCTCGAGTACTTGCACTGCCAGAGATCGTCCACAGGGAC 825

Db 853 GGTGCAGAGATTGCTCGGGCTTGTAGTACTTGCACTCGGGGACGTGGTATACCGGAC 912

QY 826 ATCAAGCCATCCAACTGCTCTCGGGGATGATGGGCACTGAGATCGCGGCTTTGGC 885

Db 913 ATCAAGCTGGAACCTCATGCTCGGACAAAGTGGCCACATCAAGATCACTGACTTTGGC 972

QY 886 GTCAAGCAACAGTTTGAGGGGACGACGCTGCTCCAGCAGCGGGGGAACCCAGCA 945

Db 973 CTCTGCAAGAGGGCATCAGTGACGGGGCCACCATGAACCTTCTGTGGGACCCCGAG 1032

QY 946 TTCTAGCCCGCCGAGGCCATTCTGATTCCGCCAGAGCTTCAGTGGGAGGCCTTGGAT 1005
Db 1033 TACCTGGCGCTGAGGTGCTGAGGACAATGACTA-----TGGCCGGCGCTGGAC 1083
QY 1006 GTATGGGCACTGGCGTCACGTTGTACTGCTTTGTCTATGGGAAGTGCCCATTCATCGAC 1065
Db 1084 TGGTGGGGCTGGGTGTGGTCACTACGAGATGATGTCGGCGCGCTGCCCTTCTACAAC 1143
QY 1066 GATTTCATCTGGCCCTCCACAGGAAGATCAAGAAATGAGCCCGTGGTGTTCCTGAGGAG 1125
Db 1144 CAGGACCACGAGCGCTCTTCGAGCTCATCTCATGGAAGAGATCCGCTTCCC-----G 1197
QY 1126 CCAGAAATCAGGAGGAGCTCAAGACCTGATCCTGAGATGTTAGACAAGATCCCGAG 1185
Db 1198 CGCAGGCTCAGCCCGAGGCCAAGTCCCTGCTTGTGGGCTGCTTTAAGAGAGACCCCAAG 1257
QY 1186 ACGAATTTGGGTG 1200
Db 1258 CAGAGGCTTGGTGG 1272

RESULT 13

US-09-023-655-1004
; Sequence 1004, Application US/09023655
; Patent No. 6607879

GENERAL INFORMATION:

APPLICANT: Cocks, Benjamin G.

APPLICANT: Susan G. Stuart

APPLICANT: Jeffrey J. Seilhamer

TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE

TITLE OF INVENTION: EXPRESSION

NUMBER OF SEQUENCES: 1508

CORRESPONDENCE ADDRESS:

ADDRESSEE: INCYTE PHARMACEUTICALS, INC.

STREET: 3174 PORTER DRIVE

CITY: PALO ALTO

STATE: CALIFORNIA

COUNTRY: USA

ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2

CURRENT APPLICATION DATA: US/09/023,655

APPLICATION NUMBER: HEREWITH

CLASSIFICATION:

PRIOR APPLICATION DATA:

FILING DATE:

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Zeller, Karen J.

REGISTRATION NUMBER: 37,071

REFERENCE/DOCKET NUMBER: PA-0001 US

TELECOMMUNICATION INFORMATION:

TELEPHONE: (650) 855-0555

TELEFAX: (650) 845-4166

INFORMATION FOR SEQ ID NO: 1004:

SEQUENCE CHARACTERISTICS:

LENGTH: 1599 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

IMMEDIATE SOURCE:

LIBRARY: GENBANK

CLONE: g178325

US-09-023-655-1004

Query Match 5.5%; Score 83.4; DB 4; Length 1599;

Best Local Similarity 51.3%; Pred. No. 6.7e-12;

Matches 254; Conservative 0; Mismatches 226; Indels 15; Gaps 2;

QY 706 GGGCCCGTCATGGAAGTGCCTGTGACAAGCCCTTCTCGAGGAGCAAGCTCGCCTCTAC 765
DB 793 GAGCTGTTCTTCCACCTGTCCGGAGCGGTCTTTCACAGAGAGCGGCCCGGTTTAT 852
QY 766 CTGCGGAGCTATCCTCGGGCTCGAGTACTTGCACTTGGCACTGCCAAGATCGTCCACAGGAC 825
DB 853 GGTGCAGAGATTGTCTCGGCTCTTGAGTACTTGCACTCGCGGAGCGTGGTATACCGGAC 912
QY 826 ATCAAGCCATCAACCTCTCTCGGGGATGATGGGACGTTGAAGATGCCGACTTTGGC 885
DB 913 ATCAAGCTGGAACCTCATGCTGACAAAGATGGCCACATCAAGATCACTGACTTTGGC 972
QY 886 GTCAAGCAACAGTTTGGAGGGAACGACGCTCAGCTGTCCAGCACGCGCGGAAACCCAGCA 945
DB 973 CTCTGCAAGAGGGATCATGTCACCGGGCCACCATGAAACCTTCTGTGGACCCCGAG 1032
QY 946 TTCAATGGCCCCGAGGCCATTTCTGATTCGCGCCAGAGCTTCAGTGGGAAGGCCCTTGGAT 1005
DB 1033 TACCTGGGCGCTGAGTGTGGTGTGATGACGAGTATGATGCGGCGGCTGCCCTTCTACAC 1083
QY 1006 GTATGGGCCACTGGCTCAGCTTGTACTGCTTGTCTATGGGAAGTGGCCATTCATCGAC 1065
DB 1084 TGGTGGGGCTGGTGTGGTGTGATGACGAGTATGATGCGGCGGCTGCCCTTCTACAC 1143
QY 1066 GATTTTCATCTCGGCCCTCCACAGGAAGATCAAGAATGAGCCCGTGTGTCTCTGAGGAG 1125
DB 1144 CAGGACCAGAGCGCTCTCTCAGCTCATCTCTCATGGAAGATCGGCTTCCC-----G 1197
QY 1126 CCAGAAATCAGGAGAGCTCAAGACCTGATCTCTGAGATGTTAGACAGAAATCCCGAG 1185
DB 1198 CGCAGCTCAGCCCGAGGCCAAGTCCCTGTTGCTGTGGCTGCTTAAGAAGGACCCCAAG 1257
QY 1186 ACGAATTTGGGGTG 1200
DB 1258 CAGAGCTTGGTGGG 1272

RESULT 14

US-09-799-875-15
; Sequence 15, Application US/09799875
; Patent No. 6638721
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachel
; APPLICANT: Kapeller-Libermann, Rosana
; APPLICANT: Williamson, Mark
; TITLE OF INVENTION: No. 6638721el Human Protein Kinases and Uses
; FILE REFERENCE: 35800/209996
; CURRENT APPLICATION NUMBER: US/09/799,875
; PRIOR FILING DATE: 2001-03-06
; PRIOR APPLICATION NUMBER: 60/182,059
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: 09/659,287
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 1257
; TYPE: DNA
; ORGANISM: Homo sapiens

US-09-799-875-15

Query Match 5.5%; Score 82.8; DB 4; Length 1257;
Best Local Similarity 53.1%; Pred. No. 8.7e-12;
Matches 224; Conservative 0; Mismatches 192; Indels 6; Gaps 2;

QY 586 GTGTACAGGAGATTGCCATCTCTGAAGAAGCTGGACCACTGAAATGTGGTCAAACTGATC 645
DB 211 GTCTTCGGGAGCTGAGATCTCTGAGGAGATCGACACGCTCTCTCTGTGTAACCTCTGG 270
QY 646 GAGGTCTTGATGACCCAGCTGAGGACAACTCTATTGGTGTGTTGACCTCTCTGAGAAAG 705

DB 271 TACTCTTCCAGGAC--GAGGAGGACATGTTATGTTGCTGGACCTGCTACTTGGGCGG 327
QY 706 GGGCCCGTCATGGAAGTGCCTGTGACAAGCCCTTCTCGAGGAGCAAGCTCGCCTCTAC 765
DB 328 GACTGCGCTACCACCTGACGAGAACGTCAGTCTCTCGAGGACACGCTGAGGCTGTAC 387
QY 766 CTGCGGAGCTATCCTCGGGCTCGAGTACTTGCACTGCCAAGATCGTCCACAGGAC 825
DB 388 ATCTCGAGATGGCACTGGCTCTGAGTACTCTGCGCGCCAGACATCATCCACAGAGAT 447
QY 826 ATCAAGCCATCAACCTCTCTCGGGGATGATGGGACGTTGAAGATGCCGACTTTGGC 885
DB 448 GTCAAGCTGGAACATTTCTCTGATGAGAGAGGACATGCACCTCAGCGACTTCAAC 507
QY 886 GTCAAGCAACAGTTTGGAGGGAACGACGCTCAGCTGTCCAGCACGCGCGGAAACCCAGCA 945
DB 508 ATTGCCACCATCATCAAG--ACGGGAGCGGGCAGCGCATTAGCAGGACCAAGCGG 564
QY 946 TTCAATGGCCCCGAGGCCATTTCTGATTCGCGCCAGAGCTTCAGTGGGAAGGCCCTTGGAT 1005
DB 565 TACATGGCTCCGAGATCTTCCACTCTTTTGTCAACGCGCGGACCGGCTACTCTCTCGAG 624
QY 1006 GT 1007
DB 625 GT 626

RESULT 15

US-09-799-875-13
; Sequence 13, Application US/09799875
; Patent No. 6638721
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachel
; APPLICANT: Kapeller-Libermann, Rosana
; APPLICANT: Williamson, Mark
; TITLE OF INVENTION: No. 6638721el Human Protein Kinases and Uses
; FILE REFERENCE: 35800/209996
; CURRENT APPLICATION NUMBER: US/09/799,875
; PRIOR FILING DATE: 2001-03-06
; PRIOR APPLICATION NUMBER: 60/182,059
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: 09/659,287
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 1826
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (17)....(1273)
US-09-799-875-13

Query Match 5.5%; Score 82.8; DB 4; Length 1826;
Best Local Similarity 53.1%; Pred. No. 1e-11;
Matches 224; Conservative 0; Mismatches 192; Indels 6; Gaps 2;

QY 586 GTGTACAGGAGATTGCCATCTCTGAAGAAGCTGGACCACTGAAATGTGGTCAAACTGATC 645
DB 227 GTCTTCGGGAGCTGAGATCTCTGAGGAGATCGACACGCTCTCTCTGTGTAACCTCTGG 286
QY 646 GAGGTCTTGATGACCCAGCTGAGGACAACTCTATTGGTGTGTTGACCTCTCTGAGAAAG 705
DB 287 TACTCTTCCAGGAC--GAGGAGGACATGTTATGTTGCTGGACCTGCTACTTGGGCGG 343
QY 706 GGGCCCGCTCATGGAAGTGCCTGTGACAGCCCTTCTCGAGGAGCAAGCTCGCCTCTAC 765
DB 344 GACTGCGCTACCACTCGCAGCAGAACGTCGAGTTCTCCGAGGACACGCTGAGGCTGTAC 403
QY 766 CTGCGGAGCTCATCTCTGGGCGCTCGAGTACTGCACTGCGCAAGATCTCCACAGGAC 825

```

Db 404 ATCTGGAGATGGCACTGGCTCTGGACTACCTGCGCGCCAGCACATCATCCACAGAGAT 463
QY 826 ATCAAGCCATCCAACTCTCTCTGGGGGATGATGGGCACGTGAAGATCGCGACTTTGGC 885
Db 464 GTCAAGCCCTGACAAATCTCTCTGGATGAGAGAGGACATGCACACCTGACCGACTTCAAC 523
QY 886 GTCAGCAACCAAGTTTGAAGGGAACGACGCTCAGCTGTCCAGCACGCGCGGMAACCCAGCA 945
Db 524 ATTGCCACCATCATCAAGG---ACGGGAGCGGGCGACGGCATTAGCAGGCAACCAAGCCG 580
QY 946 TTCATGGCCCCGAGGGCCATTTCGATTCCGGCCAGAGCTTCAAGTGGGAAGGCCCTTGGAT 1005
Db 581 TACATGGCTCCGGAGATCTTCCACTCTTTTGTCAACGGGGGACCGGCTACTCCTTCGAG 640
QY 1006 GT 1007
Db 641 GT 642

```

Search completed: July 25, 2004, 11:37:16
 Job time : 147 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 24, 2004, 22:18:04 ; Search time 897 seconds
(without alignments)
7175.050 Million cell updates/sec

Title: US-10-690-617-1_COPY_173_1687
Perfect score: 1515
Sequence: 1 atggagggggtccagctgt.....aggaagacgaggtgcattcc 1515

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_29Jan04.*

- 1: geneseqn1980s.*
- 2: geneseqn1990s.*
- 3: geneseqn2000s.*
- 4: geneseqn2001as.*
- 5: geneseqn2001bs.*
- 6: geneseqn2002s.*
- 7: geneseqn2003as.*
- 8: geneseqn2003bs.*
- 9: geneseqn2003cs.*
- 10: geneseqn2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1515	100.0	2190	6	ABL58698 Human kin
2	1510.2	99.7	1937	6	ABD30566 Human kin
3	1507	99.5	2018	6	ABZ11551 Human pol
4	1505.4	99.4	1611	6	ABX97035 Human NOV
5	1490.6	98.4	2711	6	ABK49563 Human CDN
6	1458.8	96.3	1542	4	AAS06710 Polynucle
7	1457.2	96.2	3501	5	AAD08635 Human kin
8	1455.8	96.1	1547	6	ABX97034 Human NOV
9	1381.4	91.2	1725	6	ABX97036 Human NOV
10	1236.6	81.6	3411	9	ADB53308 Primary r
11	660.2	43.6	2545	3	Aaz29223 Human cel
12	651.2	43.0	4942	9	ADB75227 Prostate
13	651	43.0	1804	6	ABK92170 Prostate
14	651	43.0	1804	9	ADB75229 Prostate
15	643	42.4	2056	4	Aai57819 Human pol
16	641.6	42.3	4982	7	ABX34704 Human mdd
17	641.4	42.3	1767	4	Aaf44686 Novel pro
18	585.6	38.7	2268	4	Aai59606 Human pol
19	585.6	38.7	2268	4	Aai59605 Human pol
20	529.6	35.0	705	4	AAS27190 cDNA enco
21	529.6	35.0	705	9	ABK43592 DNA enco
22	529.6	35.0	705	9	ADB93368 Human CDN
23	529	34.9	731	4	Aah99185 Human pro

24	526	34.7	2411	9	ADC99145	Adc99145 Human KPP
25	518	34.2	1927	4	AAI57820	Aai57820 Human pol
26	361	23.8	29629	6	ABL58699	AbL58699 Human kin
27	354.4	23.4	432	9	ADE85632	AdE85632 Farnesyl
28	348.8	23.0	439	4	AAH99412	Aah99412 Human pro
29	272	18.0	876	6	ABK34394	Abk34394 Human CDN
30	262.6	17.3	491	8	ACH14059	Ach14059 Human adu
31	244.6	16.1	485	8	ACH15568	Ach15568 Human adu
32	240	15.8	331	3	AAC02673	Aac02673 Human sec
33	202.4	13.4	1174	4	ABL19199	AbL19199 Drosophil
34	201.2	13.3	469	4	AAI15690	Aai15690 Probe #56
35	201.2	13.3	469	4	ABA57768	AbA57768 Human foe
36	201.2	13.3	469	4	AAI37350	Aai37350 Probe #60
37	201.2	13.3	469	4	ABA27140	AbA27140 Probe #56
38	201.2	13.3	469	4	AAK31464	Aak31464 Human bon
39	201.2	13.3	469	4	AAK05833	Aak05833 Human bra
40	201.2	13.3	469	4	ABS31146	AbS31146 Human liv
41	201.2	13.3	469	6	ABS06218	AbS06218 Human gen
42	201.2	13.3	10800	4	AAK83329	Aak83329 Human imm
43	201	13.3	218	4	AAI24914	Aai24914 Probe #14
44	201	13.3	218	4	ABA70385	AbA70385 Human foe
45	201	13.3	218	4	AAI50533	Aai50533 Probe #19

ALIGNMENTS

RESULT 1

ABL58698
ID ABL58698 standard; cDNA: 2190 BP.

XX ABL58698;

DT 27-AUG-2002 (first entry)

XX Human kinase encoding cDNA.

Human; kinase protein; calcium/calmodulin-dependant protein kinase; eye;
retinoblastoma; brain; kinase modulator; gene; SNP;
single nucleotide polymorphism; ss.

OS Homo sapiens.

XX Key Location/Qualifiers

FT 5'UTR 1..172

FT CDS /tag= a

FT /tag= b

FT variation /product= "human kinase"

FT replace(1296, G)

FT /tag= c

3'UTR 1691..2190 /standard_name= "single nucleotide polymorphism"

FT /tag= d

WO200224920-A2.

28-MAR-2002.

19-SEP-2001; 2001WO-US029161.

19-SEP-2000; 2000US-0233493P.

13-NOV-2000; 2000US-0247031P.

06-DEC-2000; 2000US-00729995.

(PEXE) PE CORP NY.

Beasley EM, Wei M, Bonazzi VR, Sanders R, Di Francesco V;

WPI: 2002-404955/43.

P-PSDB; ABB83054.

Novel peptide designated as human kinase useful as target for diagnosing

a disease or predisposition to the disease mediated by the peptide.

Claim 4b; Fig 1; 89pp; English.

The invention relates to an isolated peptide designated human kinase (HK), that has homology to members of the calcium/calmodulin-dependent protein kinase kinase subfamily. The mechanism of action of the protein of the invention is that of a kinase modulator. The human kinase of the invention is useful for creating a pharmaceutical composition for treating a disease or condition mediated by the human kinase. HK is also useful to provide a target for diagnosing a disease or predisposition to disease mediated by HK, and is also useful in pharmacogenic analysis. HK is useful for treating a disorder characterised by absence of an inappropriate or unwanted expression of HK, also as an immunogen to raise antibodies by administering HK to a mammalian organism e.g. rat, rabbit or mouse. Nucleic acids of the invention are useful as hybridisation probes for mRNA, transcript/cDNA and genomic DNA. Recombinant host cells expressing a native HK are useful for assaying compounds that stimulate or inhibit HK function. Nucleic acids of the invention are also useful for producing transgenic animals. Experimental data indicates that kinase proteins of the present invention are expressed in humans in the eye (retinoblastomas) and brain. The current sequence represents cDNA encoding the human kinase of the invention

Sequence 2190 BP; 457 A; 655 C; 708 G; 370 T; 0 U; 0 Other;

Query Match 100.0%; Score 1515; DB 6; Length 2190;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1515; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 ATGAGGGGGGTTCAGCTGTCTGTCGACGATCCTCGGGCAGAGCTGGTAGACGGGTG 60
DB 173 ATGAGGGGGGTTCAGCTGTCTGTCGACGATCCTCGGGCAGAGCTGGTAGACGGGTG 232
QY 61 GCGACCATCATGATGACTACTTGGAGGAGCAGATGGTGGCCAGAGCCCTACTAGAAAC 120
DB 233 GCAGCCATCATGATGACTACTTGGAGGAGCAGATGGTGGCCAGAGCCCTACTAGAAAC 292
QY 121 GGTGTGACCCCCACACAGGGCCAGAGCTGCCTCTGTGATCCCTGGCAGTACTTCAAGA 180
DB 293 GGTGTGACCCCCACACAGGGCCAGAGCTGCCTCTGTGATCCCTGGCAGTACTTCAAGA 352
QY 181 CTGCTCCAGCCCGGCTTACGCTCTCAGCCAGGAGCTTCCCTACAGGAGCGGCCAGCA 240
DB 353 CTGCTCCAGCCCGGCTTACGCTCTCAGCCAGGAGCTTCCCTACAGGAGCGGCCAGCA 412
QY 241 GGAAGCTATCTGAGGCGAGGCTGGCCCTTATGCCACGGGGCTGCCAGCCCATCTCC 300
DB 413 GGAAGCTATCTGAGGCGAGGCTGGCCCTTATGCCACGGGGCTGCCAGCCCATCTCC 472
QY 301 CCCGGGCTTGGGGAGGCGCCACCATCGAGTCCCAACAGTGGCCATCTCAGATGCAGAG 360
DB 473 CCCGGGCTTGGGGAGGCGCCACCATCGAGTCCCAACAGTGGCCATCTCAGATGCAGAG 532
QY 361 GACTGCTGAGTGAACAGTACAAGTCAAGCTCAGAGTGCAGATTCGCAAGGGTGCCTACGGT 420
DB 533 GACTGCTGAGTGAACAGTACAAGTCAAGCTCAGAGTGCAGATTCGCAAGGGTGCCTACGGT 592
QY 421 GTGCTGAGCTGGCCCTACAGCAAGAGTGAAGCAGACATCATGCAATGAAGTCCCTTTC 480
DB 593 GTGCTGAGCTGGCCCTACAGCAAGAGTGAAGCAGACATCATGCAATGAAGTCCCTTTC 652
QY 481 AAAAGAAGTCTACTGAAGCAGTATGCTTTCACGTCGCCCTCCCGGAGGGGTCCAG 540
DB 653 AAAAGAAGTCTACTGAAGCAGTATGCTTTCACGTCGCCCTCCCGGAGGGGTCCAG 712
QY 541 GTGCCCCAGGAGGACAGCCAGCAGCTGCTGCCCTTGGAGCGGGTGTACAGGAGATT 600
DB 713 GTGCCCCAGGAGGACAGCCAGCAGCTGCTGCCCTTGGAGCGGGTGTACAGGAGATT 772
QY 601 GCATCTCAAGAGCTGACACAGCTGAATGGTCAAACTGATCAGGCTCCTGGATGAC 660
DB 773 GCATCTCAAGAGCTGACACAGCTGAATGGTCAAACTGATCAGGCTCCTGGATGAC 832
```

RESULT 2

AAD30566

ID AAD30566 standard; cDNA; 1937 BP.

XX

AC AAD30566;

XX

DT 21-MAY-2002 (first entry)

XX

DE Human kinase polypeptide (PKIN-19) cDNA.

XX

KW Human; kinase polypeptide; PKIN-19; gene therapy; Addison's disease;

KW leukaemia; immune disorder; lymphoma; melanoma; developmental disorder;
KW acquired immunodeficiency syndrome; AIDS; allergy; anaemia; hypertension;
KW asthma; Crohn's disease; rheumatoid arthritis; Bursitis; atherosclerosis;
KW cirrhosis; hepatitis; psoriasis; Cushing's syndrome; cytostatic; cancer;
KW cholestasis; cardiant; cardiovascular disorder; Niemann-Pick's disease;
KW lipid disorder; fatty liver; Gaucher's disease; myocardial infarction;
KW drug screening; transgenic animal; antiinflammatory; hepatotropic;
KW hypotensive; anti-HIV; enzyme; ss.

XX Homo sapiens.

OS
XX
FH Key Location/Qualifiers
CDS 65..1582
FT /*tag= a
FT /*product= "Human PKIN-19"

PN WO200208399-A2.

XX 31-JAN-2002.

XX 20-JUL-2001; 2001WO-US023092.

XX 21-JUL-2000; 2000US-0220038P.

PR 28-JUL-2000; 2000US-0222112P.

PR 04-AUG-2000; 2000US-0222831P.

PR 11-AUG-2000; 2000US-0224729P.

XX (INCY-) INCYTE GENOMICS INC.

PA (THOR/) THORNTON M.

PA Thornton M, Yue H, Khan FA, Gururajan R, Hafalia AJA, Walia NK;

XX Patterson C, Ramkumar J, Gandhi AR, Policky JL, Baughn MR;

XX Tribouley CM, Bandman O, Nguyen DB, Lu Y, Burford N, Lal P, Ding L;

XX Yao MG, Elliott VS, Recipon SA, Kearney L, Lu DAM, Greenwald SR;

XX Tang YT, Xu Y, Walsh RT, Gietzen KJ, Yang J, Hillman JL;

XX WPI; 2002-206083/26.

DR P-PDSB; AAE19161.

XX New human kinase polypeptide, useful in diagnosis, prevention and

PT treatment of cancer, immune disorder, growth and developmental disorder,

PT cardiovascular disorder and lipid disorder.

XX Claim 5; Page 193-194; 196pp; English.

XX The present invention relates to an isolated human kinase polypeptide
CC (PKIN) or a biologically active/immunogenic fragment of PKIN. PKIN is
CC useful for diagnosing, treating and preventing cancer (e.g., leukaemia,
CC lymphoma, melanoma), an immune disorder (e.g., acquired immunodeficiency
CC syndrome (AIDS), Addison's disease, allergy, anaemia, asthma, Crohn's
CC disease, rheumatoid arthritis), a growth and developmental disorder (e.g.
CC bursitis, cirrhosis, hepatitis, psoriasis, Cushing's syndrome), a
CC cardiovascular disorder (e.g., atherosclerosis, hypertension, myocardial
CC infarction), and a lipid disorder (e.g., fatty liver, cholestasis,
CC Gaucher's disease, Niemann-Pick's disease). PKIN is useful in a number of
CC drug screening techniques and to analyse the proteome of a tissue or cell
CC type. PKIN is useful for creating knockin humanised animals or transgenic
CC animals to model human diseases, in somatic or germline gene therapy, to
CC generate a transcript image of a tissue or cell type, for detecting
CC differences in the chromosomal location due to translocation, inversion,
CC etc., among normal, carrier or affected individuals, and as hybridisation
CC probes for mapping naturally occurring genomic sequences. PKIN is useful
CC in southern or northern analysis, dot blot or other membrane-based
CC technologies, in PCR technologies, in dipstick, pin, microformat enzyme
CC linked immunosorbent (ELISA)-like assays and in microarrays utilising
CC fluids or tissues from patients to detect altered PKIN expression. The
CC present sequence is human PKIN-19 cDNA

XX Sequence 1937 BP; 425 A; 567 C; 603 G; 342 T; 0 U; 0 Other;

XX Query Match 99.7%; Score 1510.2; DB 6; Length 1937;

XX Best Local Similarity 99.8%; Pred. No. 0;

XX Matches 1512; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ATGGAGGGGGTCCAGCTGTCTGTCTGCCAGGATCTCTCGGCGAGAGCTGGTAGAACGGGTG 60
DB 65 ATGGAGGGGGTCCAGCTGTCTGTCTGCCAGGATCTCTCGGCGAGAGCTGGTAGAACGGGTG 124
QY 61 GGAGCCATCGATGTGACTCACTTTGGAGGAGGAGATGGTGGCCAGAGCTACTAGAAAC 120
DB 125 GGAGCCATCGATGTGACTCACTTTGGAGGAGGAGATGGTGGCCAGAGCTACTAGAAAC 184
QY 121 GGTGTGACCCCCCACCACGGGCCAGAGCTGCTCTGTGTATCCCTGGCAGTACTTCAAGA 180
DB 185 GGTGTGACCCCCCACCACGGGCCAGAGCTGCTCTGTGTATCCCTGGCAGTACTTCAAGA 244
QY 181 CTGCTCCAGCCCGGCTAGCCCTCTCAGCCAGGAAGCTTTCCCTACAGGAGCGGCCAGCA 240
DB 245 CTGCTCCAGCCCGGCTAGCCCTCTCAGCCAGGAAGCTTTCCCTACAGGAGCGGCCAGCA 304
QY 241 GGAAGCTATCTGAGCGCAGAGCTGGGCCCTTATGCCACGGGGCTCCAGGCCATCTCTCC 300
DB 305 GGAAGCTATCTGAGCGCAGAGCTGGGCCCTTATGCCACGGGGCTCCAGGCCATCTCTCC 364
QY 301 CCCCGGGCTGGCGGAGGCCACCATCAGTCCCCACACGCTGGCCATCTCAGATGAGAG 360
DB 365 CCCCGGGCTGGCGGAGGCCACCATCAGTCCCCACACGCTGGCCATCTCAGATGAGAG 424
QY 361 GACTGCGTGCAGTGAACCAAGTACAAAGCTGCAGAGTGAGATTGGCAAGGGTGCCTACGGT 420
DB 425 GACTGCGTGCAGTGAACCAAGTACAAAGCTGCAGAGTGAGATTGGCAAGGGTGCCTACGGT 484
QY 421 GTGGTGAGCTGGCTTACAAAGTGAAGCAGACACTATGCAATGAAAGTCTTTTTC 480
DB 485 GTGGTGAGCTGGCTTACAAAGTGAAGCAGACACTATGCAATGAAAGTCTTTTTC 544
QY 481 AAAAAAGATTACTGAAGCAGATATGCTTTTCCACGTCGCCCTCCCCCGAGAGGGTCCACG 540
DB 545 AAAAAAGATTACTGAAGCAGATATGCTTTTCCACGTCGCCCTCCCCCGAGAGGGTCCACG 604
QY 541 GCTGCCAGGGAGGACCAAGCAGAGTGTGCTGCCCTGGAGCGGGTGTACCAGGAGATT 600
DB 605 GCTGCCAGGGAGGACCAAGCAGAGTGTGCTGCCCTGGAGCGGGTGTACCAGGAGATT 664
QY 601 GCATCTCTGAAGAAGCTGGACCAAGTGAATGTGGTCAAACTGANTCAGAGTCTCTGGATGAC 660
DB 665 GCATCTCTGAAGAAGCTGGACCAAGTGAATGTGGTCAAACTGANTCAGAGTCTCTGGATGAC 724
QY 661 CCAGCTGAGGACAACTCTATTGTGTGTTGACCTCTCAGAAAGGGGCCGCTCATGGAA 720
DB 725 CCAGCTGAGGACAACTCTATTGTGTGTTGACCTCTCAGAAAGGGGCCGCTCATGGAA 784
QY 721 GTGCCCTGTGACAAAGCCCTTCTCGGAGGAGCAAGCTCGCCTCTACCTGGGGACGTCATC 780
DB 785 GTGCCCTGTGACAAAGCCCTTCTCGGAGGAGCAAGCTCGCCTCTACCTGGGGACGTCATC 844
QY 781 CTGGGCTCGAGTACTTGCACTGCCAGAGATGCTCCAGGGACATCAAGCCATCCAAC 840
DB 845 CTGGGCTCGAGTACTTGCACTGCCAGAGATGCTCCAGGGACATCAAGCCATCCAAC 904
QY 841 CTGCTCTCTGGGGATGATCGGCACAGTGAAGATCGCCGACTTTGGCGTCAAGCAACAGTTT 900
DB 905 CTGCTCTCTGGGGATGATCGGCACAGTGAAGATCGCCGACTTTGGCGTCAAGCAACAGTTT 964
QY 901 GAGGGGAACAGCGCTCAGCTGCCAGCACGGGGGAACCCAGCAATTCATGGCCCCCGAG 960
DB 965 GAGGGGAACAGCGCTCAGCTGCCAGCACGGGGGAACCCAGCAATTCATGGCCCCCGAG 1024
QY 961 GCATTTCTGATTCGGCCAGAGCTTCAGTGGGAAGCCCTTGAATGTATGGGCCACTGGC 1020
DB 1025 GCATTTCTGATTCGGCCAGAGCTTCAGTGGGAAGCCCTTGAATGTATGGGCCACTGGC 1084
QY 1021 GTACGTTGCTACTGCTTTGCTTATCGGAAGTCCCATTCATCGAGATTTTCATCTGGCC 1080
DB 1085 GTACGTTGCTACTGCTTTGCTTATCGGAAGTCCCGCTTCATCGAGATTTTCATCTGGCC 1144

QY 1081 CTCACAGAGATCAAGATGAGCCCGTGGTGTTCCTGAGGACCCAGAAATCAGCGAG 1140
DB 1145 CTCACAGAGATCAAGATGAGCCCGTGGTGTTCCTGAGGACCCAGAAATCAGCGAG 1204
QY 1141 GAGCTCAAGGACCTGATCTGAAGATGTTAGACAAGATCCCGACGAGAAATTCGGGTG 1200
DB 1205 GAGCTCAAGGACCTGATCTGAAGATGTTAGACAAGATCCCGACGAGAAATTCGGGTG 1264
QY 1201 CCAGACATCAAGTTGACCCCTTGGGTGACCAAGAACGGGGAGGACCCCTTCTTCGGAG 1260
DB 1265 CCAGACATCAAGTTGACCCCTTGGGTGACCAAGAACGGGGAGGACCCCTTCTTCGGAG 1324
QY 1261 GAGGAGCACTGACGCTGGTGGAGTGACAGAGGGGAGGTTAAGAACTCAGTCAGGCTC 1320
DB 1325 GAGGAGCACTGACGCTGGTGGAGTGACAGAGGGGAGGTTAAGAACTCAGTCAGGCTC 1384
QY 1321 ATCCCCAGCTGGAACCAAGCTGATCTGGTGAAGTCCATGCTGAGGAAGGTTCCCTTTGGG 1380
DB 1385 ATCCCCAGCTGGAACCAAGCTGATCTGGTGAAGTCCATGCTGAGGAAGGTTCCCTTTGGG 1444
QY 1381 AACCGTTTGAAGCCCAAGGACGAGGAGGAAGCGATCATGTCTGCTCCAGGAACCTA 1440
DB 1445 AACCGTTTGAAGCCCAAGGACGAGGAGGAAGCGATCATGTCTGCTCCAGGAACCTA 1504
QY 1441 CTGCTGAAGAGGCTTTGGTGAAGGGGCGAAGCCAGAGCTCCCGGCGTCCAGGA 1500
DB 1505 CTGCTGAAGAGGCTTTGGTGAAGGGGCGAAGCCAGAGCTCCCGGCGTCCAGGA 1564
QY 1501 GACGAGGCTGCATCC 1515
DB 1565 GACGAGGCTGCATCC 1579
RESULT 3
ID ABZ11551
XX AC ABZ11551;
XX DT 20-JAN-2003 (first entry)
XX DE Human polynucleotide SEQ ID NO 433.
XX KW Human; genome mapping; gene therapy; food supplement; virus; fungus;
cell-proliferative disorder; neurodegenerative disease; bacterial;
Parkinson's disease; Alzheimer's disease; autoimmune disease;
multiple sclerosis; diabetes; genetic disorder; wound; burn; infection;
arthritis; cytostatic; immunomodulator; nootropic; neuroprotective;
antiparkinsonian; antidiabetic; immunosuppressive; dermatological;
haemostatic; vulnery; fungicide; antibacterial; virucide; protozoacide;
antiarthritic; gene; ss.
XX OS Homo sapiens.
XX PN WO200270539-A2.
XX PD 12-SEP-2002.
XX PF 05-MAR-2002; 2002WO-US005095.
XX PR 05-MAR-2001; 2001US-00799451.
XX PA (HYSE-) HYSEQ INC.
XX PI Tang YT, Zhou P, Goodrich RW, Asundi V, Zhang J, Zhao Q, Ren F;
PI Xue AJ, Yang Y, Ma Y, Yamazaki V, Chen R, Wang Z, Ghosh M;
PI Wehrman T, Wang J, Wang D, Drmanac RT;
XX DR WPI: 2002-759812/82.
XX DR P-PSDB; ABP69334.
XX PT New polynucleotides comprising sequences assembled from expressed
sequence tags (ESTs), useful for treating cell-proliferative.

PT neurodegenerative, autoimmune, genetic, myeloid or lymphoid, or platelet
or coagulation disorders.
XX Claim 1; SEQ ID NO 433; 1012pp + Sequence Listing; English.
XX The invention relates to an isolated polynucleotide (I) comprising a
nucleotide sequence selected from any of 948 sequences (ABZ11119-
ABZ12066) or their mature protein coding portion, active domain coding
protein or complementary sequences. The polynucleotides are useful for
identifying expressed genes or for physical mapping of human genome. The
encoded polypeptides (ABP6902-ABP6949) are useful as molecular weight
markers, as a food supplement, for generating antibodies, in medical
imaging, screening and diagnostic assays and for treating cell
proliferative disorders (cancer), neurodegenerative diseases (Parkinson's
or Alzheimer's disease), autoimmune diseases (multiple sclerosis,
diabetes, lupus) genetic disorders, myeloid or lymphoid disorders,
platelet or coagulation disorders, wound, burn, incision, ulcers, liver
or lung fibrosis, infections (bacterial, viral, fungal, parasitic),
arthritis, etc. Note: The sequence data for this patent did not form part
of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 2018 BP; 453 A; 600 C; 603 G; 362 T; 0 U; 0 Other;
Query Match 99.5%; Score 1507; DB 6; Length 2018;
Best Local Similarity 99.7%; Fred. No. 0;
Matches 1510; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 1 ATGAGGGGGGTCCAGCTGTCTGTCGCGAGGATCTCGGGCAGAGCTGTAGAACGGGTG 60
DB 232 ATGAGGGGGGTCCAGCTGTCTGTCGCGAGGATCTCGGGCAGAGCTGTAGAACGGGTG 291
QY 61 GCAGCATCGATGACTCACTTGGAGAGGAGATGTGGCCGAGAGCTTACTAGAAC 120
DB 292 GCAGCATCGATGACTCACTTGGAGAGGAGATGTGGCCGAGAGCTTACTAGAAC 351
QY 121 GGTGTGACCCCGACACGCGGCGAGAGCTGCTCTGTGATCCCTGGCAGTACTTCAAGA 180
DB 352 GGTGTGACCCCGACACGCGGCGAGAGCTGCTCTGTGATCCCTGGCAGTACTTCAAGA 411
QY 181 CTGCTCCAGCGCGCTAGCTCTCAGCCAGGAAGCTTTCCTACAGGAGCGGCAGCA 240
DB 412 CTGCTCCAGCGCGCTAGCTCTCAGCCAGGAAGCTTTCCTACAGGAGCGGCAGCA 471
QY 241 GGAGCTATCTGAGGCGCAGGCTGGGCTTATGCCACGGGGCTTCCAGCAGCATCTCC 300
DB 472 GGAAGCTATCTGAGGCGCAGGCTGGGCTTATGCCACGGGGCTTCCAGCAGCATCTCC 531
QY 301 CCCGGGCTGCGGAGGCGCCACCATCGAGTCCACACAGTGGCCATCTCAGATCAGAG 360
DB 532 CCCGGGCTGCGGAGGCGCCACCATCGAGTCCACACAGTGGCCATCTCAGATCAGAG 591
QY 361 GACTGCGTGAGCTGAACAGTACAGCTGAGAGTGGCAAGGGTGCCTACGGT 420
DB 592 GACTGCGTGAGCTGAACAGTACAGCTGAGAGTGGCAAGGGTGCCTACGGT 651
QY 421 GTGTGAGGCTGGCTTACAAAGTGAACAGACACTATGCAATGAAGTCTTTC 480
DB 652 GTGTGAGGCTGGCTTACAAAGTGAACAGACACTATGCAATGAAGTCTTTC 711
QY 481 AAAAAGAGTTACTGAAGCAGTATGGCTTTCACGTCGCCCTCCCGGAGGGTCCCGAG 540
DB 712 AAAAAGAGTTACTGAAGCAGTATGGCTTTCACGTCGCCCTCCCGGAGGGTCCCGAG 771
QY 541 GCTGCCAGGAGGACGAGCCAGAGCTGCTGCCCTTGGAGGGGTGTACAGAGATT 600
DB 772 GCTGCCAGGAGGACGAGCCAGAGCTGCTGCCCTTGGAGGGGTGTACAGAGATT 831
QY 601 GCCATCTGAAGAGCTGACACAGTGAATGTGGTCAAACTGATCGAGGTCTCTGATGAC 660
DB 832 GCCATCTGAAGAGCTGACACAGTGAATGTGGTCAAACTGATCGAGGTCTCTGATGAC 891
QY 661 CCAGCTGAGGACAACTCTATTGTTGTTTACCTCTCTGAGAAAGGGGCCGCTCATGAA 720

Db 892 CCAGCTGAGGACAACTCTATTGTGTTTACCTCTCTGAGAAAGGGCCCGTCATGGAA 951
Qy 721 GTGCCCTGTGACAAAGCCCTTCTCGAGGAGCAAGCTCGCCTCTACCTGCGGAGGTGATC 780
Db 952 GTGCCCTGTGACAAAGCCCTTCTCGAGGAGCAAGCTCGCCTCTACCTGCGGAGGTGATC 1011
Qy 781 CTGGCCCTCGAGTACTTGTCACTGTCAGAGATCTGTCACAGGAGATCAAGCCATCCAAAC 840
Db 1012 CTGGCCCTCGAGTACTTGTCACTGTCAGAGATCTGTCACAGGAGATCAAGCCATCCAAAC 1071
Qy 841 CTGCTCTGGGGATGATGGGCACTGAGATGCGCGACTTTGGCGTCAGCAACAGATTT 900
Db 1072 CTGCTCTGGGGATGATGGGCACTGAGATGCGCGACTTTGGCGTCAGCAACAGATTT 1131
Qy 901 GAGGGGAACGAGCTCAGCTGTCAGACGCGCGGGAACCCAGCATTCATGSCCCCGCGAG 960
Db 1132 GAGGGGAACGAGCTCAGCTGTCAGACGCGCGGGAACCCAGCATTCATGSCCCCGCGAG 1191
Qy 961 GCCATTTCTGATTCGCGCCAGAGCTTCAGTGGGAAGCCCTTGGATGTATGGGCCACTGGC 1020
Db 1192 GCCATTTCTGATTCGCGCCAGAGCTTCAGTGGGAAGCCCTTGGATGTATGGGCCACTGGC 1251
Qy 1021 GTCACTGTGTACTGCTTGTCTATGGGAAGTGCCCATTCATCGACGATTCATCTCGGCC 1080
Db 1252 GTCACTGTGTACTGCTTGTCTATGGGAAGTGCCCATTCATCGACGATTCATCTCGGCC 1311
Qy 1081 CTCACAGGAAGATCAAGATGAGCCCGTGTGTTCTGAGGAGCGAGAAATCAGCGAG 1140
Db 1312 CTCATAGGAAGATCAAGATGAGCCCGTGTGTTCTGAGGAGCGAGAAATCAGCGAG 1371
Qy 1141 GAGCTCAAGGACCTGATCCTGAAGATGTAGACAAGAAATCCGAGACGAGAAATGGGGTG 1200
Db 1372 GAGCTCAAGGACCTGATCCTGAAGATGTAGACAAGAAATCCGAGACGAGAAATGGGGTG 1431
Qy 1201 CCAGACATCAAGTGTGACCCCTTGGTGACCAAGAACGGGAGGAGCCCTTCCTTCGGAG 1260
Db 1432 CCAGACATCAAGTGTGACCCCTTGGTGACCAAGAACGGGAGGAGCCCTTCCTTCGGAG 1491
Qy 1261 GAGGAGCACTGACGCTGGTGGAGTGCACAGAGGGGAGGTTAAGAACTCAGTCAGGCTC 1320
Db 1492 GAGGAGCACTGACGCTGGTGGAGTGCACAGAGGAGGAGGTTAAGAACTCAGTCAGGCTC 1551
Qy 1321 ATCCCCAGCTGGACCAAGCTGATCCTGGTGAAGTCCATGCTGAGGAAGCGTTCCTTTGGG 1380
Db 1552 ATCCCCAGCTGGACCAAGCTGATCCTGGTGAAGTCCATGCTGAGGAAGCGTTCCTTTGGG 1611
Qy 1381 AACCCGTTGAGCCCGAGGACGAGGAGGAGGATCCATGCTGCTCCAGGAACCTA 1440
Db 1612 AACCCGTTGAGCCCGAGGACGAGGAGGAGGATCCATGCTGCTCCAGGAACCTA 1671
Qy 1441 CTGGTGAAGAAGGTTTGGTGAAGGGGCAAGAGCCAGAGCTCCCGGGGTCACAGAA 1500
Db 1672 CTGGTGAAGAAGGTTTGGTGAAGGGGCAAGAGCCAGAGCTCCCGGGGTCACAGAA 1731
Qy 1501 GACGAGCTGCATCC 1515
Db 1732 GACGAGCTGCATCC 1746

RESULT 4
ABX97035
ID ABX97035 standard; cDNA; 1611 BP.
XX
AC ABX97035;
XX
DT 20-MAY-2003 (first entry)
XX
DE Human NOV15b cDNA.
KW NOVX; cytostatic; cardiant; antiarteriosclerotic; antiasthmatic; cancer;
KW hypotensive; cardiomyopathy; bronchial asthma; gene therapy; vaccine;
KW human; gene; ss.

XX OS Homo sapiens.
XX PN W0200272757-A2.
XX PD 19-SEP-2002.
XX PF 08-MAR-2002; 2002WO-US006908.
XX PR 08-MAR-2001; 2001US-0274101P.
PR 08-MAR-2001; 2001US-0274194P.
PR 08-MAR-2001; 2001US-0274281P.
PR 08-MAR-2001; 2001US-0274322P.
PR 09-MAR-2001; 2001US-0274849P.
PR 12-MAR-2001; 2001US-0275235P.
PR 13-MAR-2001; 2001US-0275578P.
PR 13-MAR-2001; 2001US-0275579P.
PR 13-MAR-2001; 2001US-0275601P.
PR 14-MAR-2001; 2001US-0276000P.
PR 16-MAR-2001; 2001US-0276776P.
PR 19-MAR-2001; 2001US-0276994P.
PR 20-MAR-2001; 2001US-0277239P.
PR 20-MAR-2001; 2001US-0277321P.
PR 20-MAR-2001; 2001US-0277327P.
PR 21-MAR-2001; 2001US-027791P.
PR 22-MAR-2001; 2001US-0277833P.
PR 23-MAR-2001; 2001US-0278152P.
PR 26-MAR-2001; 2001US-0278949P.
PR 27-MAR-2001; 2001US-0278999P.
PR 27-MAR-2001; 2001US-0279036P.
PR 28-MAR-2001; 2001US-0279344P.
PR 30-MAR-2001; 2001US-0277338P.
PR 30-MAR-2001; 2001US-0279995P.
PR 30-MAR-2001; 2001US-0280233P.
PR 02-APR-2001; 2001US-0280802P.
PR 02-APR-2001; 2001US-0280822P.
PR 02-APR-2001; 2001US-0280900P.
PR 04-APR-2001; 2001US-0281194P.
PR 13-APR-2001; 2001US-0283675P.
PR 30-APR-2001; 2001US-0287424P.
PR 02-MAY-2001; 2001US-0288066P.
PR 03-MAY-2001; 2001US-0288342P.
PR 03-MAY-2001; 2001US-0288528P.
PR 15-MAY-2001; 2001US-0291190P.
PR 16-MAY-2001; 2001US-0291099P.
PR 16-MAY-2001; 2001US-0291240P.
PR 30-MAY-2001; 2001US-0294485P.
PR 31-MAY-2001; 2001US-0294489P.
PR 31-MAY-2001; 2001US-0294899P.
PR 18-JUN-2001; 2001US-0299027P.
PR 19-JUN-2001; 2001US-0299030P.
PR 19-JUN-2001; 2001US-0299310P.
PR 10-JUL-2001; 2001US-0304354P.
PR 31-JUL-2001; 2001US-0309198P.
PR 16-AUG-2001; 2001US-0312903P.
PR 10-SEP-2001; 2001US-0318462P.
PR 12-SEP-2001; 2001US-0318770P.
PR 27-SEP-2001; 2001US-0325430P.
PR 27-SEP-2001; 2001US-0325681P.
PR 18-OCT-2001; 2001US-0330380P.
PR 31-OCT-2001; 2001US-0335301P.
PR 14-NOV-2001; 2001US-0332172P.
PR 14-NOV-2001; 2001US-0332271P.
PR 14-NOV-2001; 2001US-0332272P.
PR 14-NOV-2001; 2001US-0333184P.
PR 14-NOV-2001; 2001US-0333272P.
PR 21-NOV-2001; 2001US-0332694P.
PR 03-DEC-2001; 2001US-0337426P.
PR 03-DEC-2001; 2001US-0338092P.
PR 04-DEC-2001; 2001US-0337185P.
PR 03-JAN-2002; 2002US-0345705P.
PR 07-MAR-2002; 2002US-00092900.
XX

(CURA-) CURAGEN CORP.

PA Padigaru M, Spytek KA, Shenoy SG, Taupier RJ, Pena CE, Li L;
XX Zerhusen BD, Gusev V, Ji W, Gorman L, Miller CE, Kekuda R;
PI Patturajan M, Gangoli E, Vernet CAM, Guo X, Tchernev V;
PI Fernandes ER, Casman SJ, Malyankar UM, Gerlach V, Liu Y, Anderson D;
PI Spaderna SK, Catterton E, Burgess C, Leite M, Zhong H, Alsobrook JP;
PI Lopley DM, Rieger DK;
XX WPI: 2002-723332/78.
DR P-PSDB; ABU65068.
XX

PT NOVX polypeptides and polynucleotides, useful for preventing or treating
PT a disorder associated with aberrant NOVX expression or activity e.g.,
PT cancer, hypertension, atherosclerosis, cardiomyopathy or bronchial
PT asthma.

XX Claim 13; Page 139; 1103pp; English.

XX This invention describes novel human NOVX polypeptides which have
CC cytotstatic, cardiant, antiarteriosclerotic, antiasthmatic and hypotensive
CC activity. Pharmaceutical compositions comprising the NOVX proteins or
CC nucleic acid molecules or NOVX antibodies are useful for preventing or
CC treating a disorder associated with aberrant NOVX expression or activity
CC e.g. cancer, hypertension, atherosclerosis, cardiomyopathy or bronchial
CC asthma. The products of the invention can be used for gene therapy or in
CC a vaccine. ABX97008-ABX97185 are cDNA fragments amplified and isolated by
CC the PCR primers and probes represented in ABX13460-ABX13462 and ABX97186-
CC ABX97593. ABX97008-ABX97185 encode the NOVX proteins described in
CC ABU65041-ABU65218
XX

SX Sequence 1611 BP; 366 A; 461 C; 497 G; 287 T; 0 U; 0 Other;

Query Match 99.4%; Score 1505.4; DB 6; Length 1611;

Best Local Similarity 99.6%; Pred. No. 0;

Matches 1509; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 ATGGAGGGGGTCCAGCTGTCTGCTGCCAGGATCTCTCGGGCAGAGCTGGTAGAAACGGGTG 60
DB 52 ATGGAGGGGGTCCAGCTGTCTGCTGCCAGGATCTCTCGGGCAGAGCTGGTAGAAACGGGTG 111
QY 61 GAGGCATCGATGACTCACTTGGAGGAGGAGATGGTGGCCAGAGCTACTAGAAAC 120
DB 112 GAGGCATCGATGACTCACTTGGAGGAGGAGATGGTGGCCAGAGCTACTAGAAAC 171
QY 121 GGTGTGACCCCAACCGGSCCAGAGCTGCTCTGATCCCTGGCAGTACTTCAAGA 180
DB 172 GGTGTGACCCCAACCGGSCCAGAGCTGCTCTGATCCCTGGCAGTACTTCAAGA 231
QY 181 CTGCTCCCAACCGGSCCTAGCTCTCAGCCAGGAAGCTTTCCCTACAGGAGGGCCAGCA 240
DB 232 CTGCTCCCAACCGGSCCTAGCTCTCAGCCAGGAAGCTTTCCCTACAGGAGGGCCAGCA 291
QY 241 GGAAGTATCTGGAGCGCAGGCTGGGCCCTTATGCCACGGGGCTGCCAGGCCATCTCC 300
DB 292 GGAAGTATCTGGAGCGCAGGCTGGGCCCTTATGCCACGGGGCTGCCAGGCCATCTCC 351
QY 301 CCCCGGGCTGGGAGGCGCCACCATCGAGTCCCAACGCTGGCCATCTCAGATGAGAG 360
DB 352 CCCCGGGCTGGGAGGCGCCACCATCGAGTCCCAACGCTGGCCATCTCAGATGAGAG 411
QY 361 GACTCGTGCAGCTGAACAGTACAGCTGCAGCTGCAGTGAAGTGGCAAGGGTGCCTACGGT 420
DB 412 GACTCGTGCAGCTGAACAGTACAGCTGCAGTGAAGTGGCAAGGGTGCCTACGGT 471
QY 421 GTGGTGAGGCTGGCTACAAAGTGAAGCAGACACTATGCAATGAAGTCTTTCC 480
DB 472 GTGGTGAGGCTGGCTACAAAGTGAAGCAGACACTATGCAATGAAGTCTTTCC 531
QY 481 AAAAAGAGTTACTGAAGCAGTATGGCTTTCCACGTCGCGCTCCCGGAGAGGGTCCAG 540
DB 532 AAAAAGAGTTACTGAAGCAGTATGGCTTTCCACGTCGCGCTCCCGGAGAGGGTCCAG 591

RESULT 5

ABK49563

ID ABK49563 standard; DNA; 2711 BP.

XX

QY	661	GCCATTTCTGATTC	CGGCCACAGCTTC	CAGTGGGAGGCTT	TGGATGATGGGCCACTGGC	1028
Db	1158	GCCATTTTCTGATTC	CGGCCACAGCTTC	CAGTGGGAGGCTT	TGGATGATGGGCCACTGGC	1217
QY	1021	GTCCAGTTGTACTG	CTTTTGTTCTAT	TGGGAAGTGCCCATTC	ATCGACGATTTTCAATCCTGGCC	1080
Db	1218	GTCCAGTTGTACTG	CTTTTGTTCTAT	TGGGAAGTGCCCGTTC	ATCGACGATTTTCAATCCTGGCC	1277
QY	1081	CTCCACAGGAAGAT	CAAGATGAGCCCGT	TGTTTCTGTAGGAGCCAGAAAT	CAGCGAG	1140
Db	1278	CTCCACAGGAAGAT	CAAGATGAGCCCGT	TGTTTCTGTAGGAGCCAGAAAT	CAGCGAG	1337
QY	1141	GAGCTCAAGGACCT	GTATCCTGAAATG	TTAGACAAGAAATCCCGACGAGAAAT	TGGGGTG	1200
Db	1338	GAGCTCAAGGACCT	GTATCCTGAAATG	TTAGACAAGAAATCCCGACGAGAAAT	TGGGGTG	1397
QY	1201	CCAGACATCAAGTT	GTGACCCCTTGCGT	TGGTGACCAAGAACGGGGAGAGCCCTT	TCCTTCGGAG	1260
Db	1398	CCAGACATCAAGTT	GTGACCCCTTGCGT	TGGTGACCAAGAACGGGGAGAGCCCTT	TCCTTCGGAG	1457
QY	1261	GAGGAGCACTG	CAGCGTGTGGAGGT	GACAGAGGGGAGGTTAAGAACT	CAGTCAGGCTC	1320
Db	1458	GAGGAGCACTG	CAGCGTGTGGAGGT	GACAGAGGGGAGGTTAAGAACT	CAGTCAGGCTC	1517
QY	1321	ATCCCCAGCTGG	ACACAGTGATCCT	TGGTGAAGTCCATGCTGAGGAAGCGTT	TCCTTTGGG	1380
Db	1518	ATCCCCAGCTGG	ACACAGTGATCCT	TGGTGAAGTCCATGCTGAGGAAGCGTT	TCCTTTGGG	1577
QY	1381	AACCCGTTTGAG	CCCCCAGGCACG	GAGGGAAGAGCGATCCATGCTGCTCCAGGAAACCTA	1440	
Db	1578	AACCCGTTTGAG	CCCCCAGGCACG	GAGGGAAGAGCGATCCATGCTGCTCCAGGAAACCTA	1637	
QY	1441	CTGCTGAAGAAG	AGGTTTGGTGA	AGGGGGCAAGACCCAGAGCTCCCGGCGTCCAGGAA	1500	
Db	1638	CTGCTGAAGAAG	AGGTTTGGTGA	AGGGGGCAAGACCCAGAGCTCCCGGCGTCCAGGCT	1697	
QY	1501	GACGA	1505			
Db	1698	TACCA	1702			
RESULT 6						
AAS06710						
ID	AAS06710 standard; cDNA; 1542 BP.					
XX	AC					
XX	AAS06710;					
DT	12-SEP-2001 (first entry)					
XX	Polynucleotide sequence encoding human protein kinase #10.					
DE	Human; protein kinase; PTK; STK; cancer; cardiovascular disease;					
KW	metabolic disorder; immune related disease; neurological disorder;					
KW	neurodegenerative disorder; inflammatory disorder; infectious disease;					
XX	reproductive disorder; gene therapy; ss.					
OS	Homo sapiens.					
PN	WO200138503-A2.					
XX	31-MAY-2001.					
PD	22-NOV-2000; 2000WO-US032085.					
PP	24-NOV-1999; 99US-0167482P.					
XX	(SUGE-) SUGEN INC.					
XX	Flowman GD, Whyte D, Manning G, Sudarsanam S, Martinez R;					
PI	Flanagan P, Clary D;					
XX	WPI; 2001-343950/36.					

Db 65 ATGAGGGGGTCCAGCTGCTGCTGCCAGGATCCTCGGGCAGAGCTGCTAGAAACGGGTG 124
QY 61 CGAGCCATGATGACTCAGTTGAGGAGGAGATGGTGGCCAGAGCCTACTAGAAAC 120
Db 125 CGAGCCATGATGACTCAGTTGAGGAGGAGATGGTGGCCAGAGCCTACTAGAAAC 184
QY 121 GGTGTGGACCCCCACACCGGGCCAGAGTGCCTCTGTGATCCCTGGCAGTACTTCAAGA 180
Db 185 GGTGTGGACCCCCACACCGGGCCAGAGTGCCTCTGTGATCCCTGGCAGTACTTCAAGA 244
QY 181 CTGTCCCGACCGCGCTTAGCCTCTCAGCCAGGAAGCTTTTCCTACAGGAGCGCCAGCA 240
Db 245 CTGTCCCGACCGCGCTTAGCCTCTCAGCCAGGAAGCTTTTCCTACAGGAGCGCCAGCA 304
QY 241 GGAAGCTATCTGAGGCGCAGCTGGCCTTAGCCTTAGCAGGGGCGCTGCAGCCACATCTCC 300
Db 305 GGAAGCTATCTGAGGCGCAGCTGGCCTTAGCCTTAGCAGGGGCGCTGCAGCCACATCTCC 364
QY 301 CCCCGGCTGTGGGAGGCCACCATCGAGTCCCAACAGCTGGCCATCTCAGATGACAGAG 360
Db 365 CCCCGGCTGTGGGAGGCCACCATCGAGTCCCAACAGCTGGCCATCTCAGATGACAGAG 424
QY 361 GACTGGTGACGTGAACAGATCAAGCTGCGAGTGAATTTGGCA----- 406
Db 425 GACTGGTGACGTGAACAGATCAAGCTGCGAGTGAATTTGGCAAGGTGGGGCTGACT 484
QY 407 -----AGGTGCCCTACGCTGCTGGTGAGGCTGGCCCTACAACGAAGTGAAGAC 453
Db 485 GATGCTATCTGAGGGTCCCTACCGTGGTGGTGGCTGGCCCTACAACGAAGTGAAGAC 544
QY 454 AGACACTATGCAATGAAAGTCCCTTTCCAAAGAAAGTTACTGAAGCAGTATGGCTTTCCA 513
Db 545 AGACACTATGCAATGAAAGTCCCTTTCCAAAGAAAGTTACTGAAGCAGTATGGCTTTCCA 604
QY 514 GGTGCGCTCCCGCAGAGGTTCCAGAGTGCCTCCAGGGAGGACCAAGCAAGCAGTGTG 573
Db 605 GGTGCGCTCCCGCAGAGGTTCCAGAGTGCCTCCAGGGAGGACCAAGCAAGCAGTGTG 664
QY 574 CCCCTGGAGCGGTGTACAGAGATTCCTATCCTGAGAGAGTCTGAAGCAGTATGGCTTTCCA 633
Db 665 CCCCTGGAGCGGTGTACAGAGATTCCTATCCTGAGAGAGTCTGAAGCAGTATGGCTTTCCA 724
QY 634 GTCAAACCTGATCAGGTCCTGATGACCCAGCTGAGGACCAACTCTATTTGGTTTGAC 693
Db 725 GTCAAACCTGATCAGGTCCTGATGACCCAGCTGAGGACCAACTCTATTTGG---TTGAC 781
QY 694 CTCCTGAGAAAGGGCCCGTCAATGGAAGTGCCTGTGACAGCCCTTCTCGAGAGGACAA 753
Db 782 CTCCTGAGAAAGGGCCCGTCAATGGAAGTGCCTGTGACAGCCCTTCTCGAGAGGACAA 841
QY 754 GCTCGCCTCTACCTCGGGGAGTCAATCCTGGGCTTCGAGTACTTGCATGCCAGAGATC 813
Db 842 GCTCGCCTCTACCTCGGGGAGTCAATCCTGGGCTTCGAGTACTTGCATGCCAGAGATC 901
QY 814 GTCCACAGGACATCAAGCCATCCAACTGCTCTGCGGGATGATGGCCACGTGAAGATC 873
Db 902 GTCCACAGGACATCAAGCCATCCAACTGCTCTGCGGGATGATGGCCACGTGAAGATC 961
QY 874 GCGGACTTTGGCTCAGCAACAGTTTGGGGAGCAGCGCTCAGCTGTCAGACAGCGC 933
Db 962 GCGGACTTTGGCTCAGCAACAGTTTGGGGAGCAGCGCTCAGCTGTCAGACAGCGC 1021
QY 934 GGAACCCAGCATTCATGCCCGCCGAGGCCATTTCTGATTCGGGCCAGAGCTTCAGTGGG 993
Db 1022 GGAACCCAGCATTCATGCCCGCCGAGGCCATTTCTGATTCGGGCCAGAGCTTCAGTGGG 1081
QY 994 AAGGCTTGGATGATGGGCCACTGGGCTCAGTTGTACTGCTTTGTCTATGGGAAGTC 1053
Db 1082 AAGGCTTGGATGATGGGCCACTGGGCTCAGTTGTACTGCTTTGTCTATGGGAAGTC 1141
QY 1054 CCATTTCACGAGTTTTCATCTGGGCCCTCCACAGGAAGATCAAGAAATGAGCCCGTGGT 1113
Db 1142 CCGTTTCACGAGTTTTCATCTGGGCCCTCCACAGGAAGATCAAGAAATGAGCCCGTGGT 1201

QY 1114 TTTCTGAGGAGCCAGAAATCAGCGAGGAGCTCAAGGACCTGATCTCTGAAGATGTTAGAC 1173
Db 1202 TTTCTGAGGAGCCAGAAATCAGCGAGGAGCTCAAGGACCTGATCTCTGAAGATGTTAGAC 1261
QY 1174 AAGAATCCCGACGAGGATTCGGGTGCCAGACATCAAGTTGCACCCCTTGGGTGACCAAG 1233
Db 1262 AAGAATCCCGACGAGGATTCGGGTGCCAGACATCAAGTTGCACCCCTTGGGTGACCAAG 1321
QY 1234 AACGGGAGGAGCCCTTCTTCGAGGAGGAGCACTGCAGGCTGGTGGAGGTGACAGAG 1293
Db 1322 AACGGGAGGAGCCCTTCTTCGAGGAGGAGCACTGCAGGCTGGTGGAGGTGACAGAG 1381
QY 1294 GGGGAGGTTAAGAACTCAGTCAAGGCTCATCCCAAGTGGACCAAGTGTGTTGTTGAAG 1353
Db 1382 GAGGAGGTTAAGAACTCAGTCAAGGCTCATCCCAAGTGGACCAAGTGTGTTGTTGAAG 1441
QY 1354 TCCATGCTGAGGAAGCGTTCCCTTTGGGAACCCGTTTGGAGCCCAAGGAGGAAGAG 1413
Db 1442 TCCATGCTGAGGAAGCGTTCCCTTTGGGAACCCGTTTGGAGCCCAAGGAGGAAGAG 1501
QY 1414 CCATCCATGTTGCTCCAGGAAACCTACTGTTGAAAGAGGTTTGGTGAAGGGGCAAG 1473
Db 1502 CCATCCATGTTGCTCCAGGAAACCTACTGTTGAAAGAGGTTTGGTGAAGGGGCAAG 1561
QY 1474 AGCCAGAGCTCCCGGCGTCCAGGAACAGGAGCTGCATCC 1515
Db 1562 AGCCAGAGCTCCCGGCGTCCAGGAACAGGAGCTGCATCC 1603
RESULT 8
ABX97034
ID ABX97034 standard; cDNA; 1547 BP.
XX
AC ABX97034;
XX
DT 20-MAY-2003 (first entry)
XX
DE Human NOV15a cDNA.
XX
KW NOX; cytostatic; cardiant; antiarteriosclerotic; antiasthmatic; cancer;
KW hypotensive; cardiomyopathy; bronchial asthma; gene therapy; vaccine;
KW human; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200272757-A2.
XX
PD 19-SEP-2002.
XX
PF 08-MAR-2002; 2002WO-US006908.
XX
PR 08-MAR-2001; 2001US-0274101P.
PR 08-MAR-2001; 2001US-0274194P.
PR 08-MAR-2001; 2001US-0274281P.
PR 08-MAR-2001; 2001US-0274322P.
PR 09-MAR-2001; 2001US-0274849P.
PR 12-MAR-2001; 2001US-0275235P.
PR 13-MAR-2001; 2001US-0275578P.
PR 13-MAR-2001; 2001US-0275579P.
PR 13-MAR-2001; 2001US-0275601P.
PR 14-MAR-2001; 2001US-0276000P.
PR 16-MAR-2001; 2001US-0276776P.
PR 19-MAR-2001; 2001US-0276994P.
PR 20-MAR-2001; 2001US-0277239P.
PR 20-MAR-2001; 2001US-0277321P.
PR 20-MAR-2001; 2001US-0277327P.
PR 21-MAR-2001; 2001US-0277791P.
PR 22-MAR-2001; 2001US-0277833P.
PR 23-MAR-2001; 2001US-0278152P.
PR 26-MAR-2001; 2001US-0278894P.
PR 27-MAR-2001; 2001US-0278999P.
PR 27-MAR-2001; 2001US-0279036P.

PR 28-MAR-2001; 2001US-0279344P.
PR 30-MAR-2001; 2001US-0277338P.
PR 30-MAR-2001; 2001US-0279959P.
PR 30-MAR-2001; 2001US-0280233P.
PR 02-APR-2001; 2001US-0280802P.
PR 02-APR-2001; 2001US-0280822P.
PR 02-APR-2001; 2001US-0280900P.
PR 13-APR-2001; 2001US-0281194P.
PR 03-APR-2001; 2001US-0283675P.
PR 02-APR-2001; 2001US-0287424P.
PR 02-MAY-2001; 2001US-0288066P.
PR 03-MAY-2001; 2001US-0288342P.
PR 03-MAY-2001; 2001US-0288528P.
PR 15-MAY-2001; 2001US-0291190P.
PR 16-MAY-2001; 2001US-0291099P.
PR 16-MAY-2001; 2001US-0291240P.
PR 30-MAY-2001; 2001US-0294485P.
PR 31-MAY-2001; 2001US-0294889P.
PR 31-MAY-2001; 2001US-0294899P.
PR 18-JUN-2001; 2001US-0299027P.
PR 19-JUN-2001; 2001US-0299303P.
PR 19-JUN-2001; 2001US-0299310P.
PR 10-JUL-2001; 2001US-0304354P.
PR 31-JUL-2001; 2001US-0309198P.
PR 16-AUG-2001; 2001US-0312903P.
PR 10-SEP-2001; 2001US-0318462P.
PR 12-SEP-2001; 2001US-0318770P.
PR 27-SEP-2001; 2001US-0325430P.
PR 27-SEP-2001; 2001US-0325681P.
PR 18-OCT-2001; 2001US-0330380P.
PR 31-OCT-2001; 2001US-0335301P.
PR 14-NOV-2001; 2001US-0332172P.
PR 14-NOV-2001; 2001US-0332271P.
PR 14-NOV-2001; 2001US-0332272P.
PR 14-NOV-2001; 2001US-0333184P.
PR 14-NOV-2001; 2001US-0333272P.
PR 21-NOV-2001; 2001US-0332094P.
PR 03-DEC-2001; 2001US-0337426P.
PR 03-DEC-2001; 2001US-0338092P.
PR 04-DEC-2001; 2001US-0337185P.
PR 03-JAN-2002; 2002US-0345705P.
PR 07-MAR-2002; 2002US-00092900.
XX
XX (CURA-) CURAGEN CORP.
XX Padigar M, Spytek KA, Shenoy SG, Taupier RJ, Pena CEA, Li L;
XX Zerhusen BD, Gusev V, Ji W, Gorman L, Miller CE, Kekuda R;
PI Patturajan M, Gangoli E, Vernet CAM, Guo X, Tchernev V;
PI Fernandes ER, Casman SJ, Malyankar UM, Gerlach V, Liu Y, Anderson D;
PI Spaderna SK, Catterton E, Burgess C, Leite M, Zhong H, Alsobrook JP;
PI Lepley DM, Rieger DK;
XX WPI; 2002-723332/78.
DR P-PSDB; ABU65067.
XX
XX NOVX polypeptides and polynucleotides, useful for preventing or treating
PT a disorder associated with aberrant NOVX expression or activity e.g.,
PT cancer, hypertension, atherosclerosis, cardiomyopathy or bronchial
PT asthma.
XX
PS Claim 13; Page 138; 1103pp; English.
PS
XX This invention describes novel human NOVX polypeptides which have
CC cytostatic, cardiac, antiarteriosclerotic, antiasthmatic and hypotensive
CC activity. Pharmaceutical compositions comprising the NOVX proteins or
CC nucleic acid molecules or NOVX antibodies are useful for preventing or
CC treating a disorder associated with aberrant NOVX expression or activity
CC e.g. cancer, hypertension, atherosclerosis, cardiomyopathy or bronchial
CC asthma. The products of the invention can be used for gene therapy or in
CC a vaccine. ABX97008-ABX97185 are cDNA fragments amplified and isolated by
CC the PCR primers and probes represented in ABX13460-ABX13462 and ABX97186-
CC ABX97593. ABX97008-ABX97185 encode the NOVX proteins described in
CC ABU65041-ABU65218

XX
SQ Sequence 1547 BP; 352 A; 441 C; 474 G; 280 T; 0 U; 0 Other;
Query Match 96.1%; Score 1455.8; DB 6; Length 1547;
Best Local Similarity 98.5%; Pred. No. 0; Mismatches 17; Indels 6; Gaps 2;
Matches 1492; Conservative 0;
QY 1 ATGGAGGGGGTCCAGCTGTCTGCTCCAGGATCTCTCGGCGAGAGTGTGTAGAACGGGTG 60
DB |||||
QY 20 ATGGAGGGGGTCCAGCTGTCTGCTCCAGGATCTCTCGGCGAGAGTGTGTAGAACGGGTG 79
DB |||||
QY 61 GCAGCCATCATGTGACTACATTGGAGAGGAGATGGTGGCCAGAGCTATAGAAC 120
DB |||||
QY 80 GCAGCCATCATGTGACTACATTGGAGAGGAGATGGTGGCCAGAGCTATAGAAC 139
DB |||||
QY 121 GGTGTGAGCCCCCACCAGGGCCAGAGCTGCTCTGTGATCCCTGGCAGTACTCAAGA 180
DB |||||
QY 140 GGTGTGAGCCCCCACCAGGGCCAGAGCTGCTCTGTGATCCCTGGCAGTACTCAAGA 199
DB |||||
QY 181 CTGCTCCAGCCCGGCTAGCCTCTCAGCCAGGAAGCTTTCCCTACAGGAGCGCCAGCA 240
DB |||||
QY 200 CTGCTCCAGCCCGGCTAGCCTCTCAGCCAGGAAGCTTTCCCTACAGGAGCGCCAGCA 259
DB |||||
QY 241 GGAAGCTATCTGAGGCGCAGGCTGGGCTTATGCGACGGGGCTGCGACGACATCTCC 300
DB |||||
QY 260 GGAAGCTATCTGAGGCGCAGGCTGGGCTTATGCGACGGGGCTGCGACGACATCTCC 319
DB |||||
QY 301 CCCCAGGCTGCGGAGGCGCCACCAGTGCAGTCCACACAGTGGGCCATCTCAGATGAGAG 360
DB |||||
QY 320 CCCCAGGCTGCGGAGGCGCCACCAGTGCAGTCCACACAGTGGGCCATCTCAGATGAGAG 379
DB |||||
QY 361 GACTGCGTGCAGTGAACAGGTACAAGTGCAGAGTGAAGTGGCAAGGGTGGCTACGGT 420
DB |||||
QY 380 GACTGCGTGCAGTGAACAGGTACAAGTGCAGAGTGAAGTGGCAAGGGTGGCTACGGT 439
DB |||||
QY 421 GTGCTGAGGCTGCGCTACAAAGTGAAGAGAGAGAGTATGCAATGAAGTCTCTTTC 480
DB |||||
QY 440 GTGCTGAGGCTGCGCTACAAAGTGAAGAGAGAGTATGCAATGAAGTCTCTTTC 499
DB |||||
QY 481 AAAAGAAGTTACTGAAGCAGTATGCTTTCCAGTGCCTCCCGGAGAGGGTCCAG 540
DB |||||
QY 500 AAAAGAAGTTACTGAAGCAGTATGCTTTCCAGTGCCTCCCGGAGAGGGTCCAG 559
DB |||||
QY 541 GCTGCCAGGAGGAGCCAGCAGAGCTGCTCCCTGGAGCGGTGTACAGGAGATT 600
DB |||||
QY 560 GCTGCCAGGAGGAGCCAGCAGAGCTGCTCCCTGGAGCGGTGTACAGGAGATT 619
DB |||||
QY 601 GCATCTGAGAGAGCTGAGCAGCAGTGAATGTGGTCAAACTGATCGAGTCTCTGGATGAC 660
DB |||||
QY 620 GCATCTGAGAGAGCTGAGCAGCAGTGAATGTGGTCAAACTGATCGAGTCTCTGGATGAC 679
DB |||||
QY 661 CCAGCTGAGGAGCAACCTCTATTGTTGTTGACCTCTCGAAGAGGGGCGGCTCATGGAA 720
DB |||||
QY 680 CCAGCTGAGGAGCAACCTCTATTGTTGTTGACCTCTCGAAGAGGGGCGGCTCATGGAA 736
DB |||||
QY 721 GTCCTCTGTGACAAGCCCTTCTCGAGGAGCAAGCTCGCTCTACCTGCGGAGCGTCATC 780
DB |||||
QY 737 GTCCTCTGTGACAAGCCCTTCTCGAGGAGCAAGCTCGCTCTACCTGCGGAGCGTCATC 796
DB |||||
QY 781 CTGGGCTCGAGTACTGTCATCTGCCAGAGAGATGTCACAGGAGACATCAAGCCATCAAC 840
DB |||||
QY 797 CTGGGCTCGAGTACTGTCATCTGCCAGAGAGATGTCACAGGAGACATCAAGCCATCAAC 856
DB |||||
QY 841 CTGCTCTCGGGGATGATGGGACGTAAGATCGCGACTTTTGGCGTACGAACACAGTTT 900
DB |||||
QY 857 CTGCTCTCGGGGATGATGGGACGTAAGATCGCGACTTTTGGCGTACGAACACAGTTT 916
DB |||||
QY 901 GAGGGAAACGACGCTCAGCTGTCCAGCAGCGGGGAAACCCAGCATTATGGCCCCCGAG 960
DB |||||
QY 917 GAGGGAAACGACGCTCAGCTGTCCAGCAGCGGGGAAACCCAGCATTATGGCCCCCGAG 976
DB |||||
QY 961 GCATTCTGATTCCCGCAGAGCTTCAGTGGGAAGGCGCTTGGATGTATGGGCACTGGC 1020
DB |||||

Db 977 GCCATTTCTGATTCGGCCAGAGCTTCACTGGGAAG---TTGGATGATATGGGCCACTGGC 1033
QY 1021 GTCAAGTTCTACTGCTTCTCTATGGAAGTGCCTCATTCAGACGATTTCTATCTCTGGCC 1080
Db 1034 GTCAAGTTCTACTGCTTCTCTATGGAAGTGCCTCATTCAGACGATTTCTATCTCTGGCC 1093
QY 1081 CTCCACAGGAAGATCAAGAATGAGCCCGTGGTGTTCCTTGAGGAGCCAGAAATCAGCGAG 1140
Db 1094 CTCCACAGGAAGATCAAGAATGAGCCCGTGGTGTTCCTTGAGGAGCCAGAAATCAGCGAG 1153
QY 1141 GAGCTCAAGGACTGATCTCTGAAGATGTTAGACAAGATCCCGACGAGAAATGGGGTG 1200
Db 1154 GAGCTCAAGGACTGATCTCTGAAGATGTTAGACAAGATCCCGACGAGAAATGGGGTG 1213
QY 1201 CCAGACATCAAGTTGACCCCTTGGGTGACCAAGACGGGGAGGCCCTTCTCTCGGAG 1260
Db 1214 CCAGACATCAAGTTGACCCCTTGGGTGACCAAGACGGGGAGGCCCTTCTCTCGGAG 1273
QY 1261 GAGGAGCACTGCAGCGTGTGTGAGGTGACAGAGGGGGAGGTTAAGAACTCAGTCAGGCTC 1320
Db 1274 GAGGAGCACTGCAGCGTGTGTGAGGTGACAGAGGGGGAGGTTAAGAACTCAGTCAGGCTC 1333
QY 1321 ATCCCCAGCTGACACCGTGTATCTCTGTTGAAGTCACTGCTGAGGAGGCTTCCTTTGGG 1380
Db 1334 ATCCCCAGCTGACACCGTGTATCTCTGTTGAAGTCACTGCTGAGGAGGCTTCCTTTGGG 1393
QY 1381 AACCCGTTTGAAGAGGCTTGGTGAAGGGGCAAGAGCCAGAGCTCCCGGCGTCCAGGAA 1513
Db 1501 GACGAGGCTGCATCC 1515
Db 1514 GACGAGGCTGCATCC 1528

RESULT 9
ABX97036

ID ABX97036 standard; cDNA; 1725 BP.
XX
AC ABX97036;
XX
DT 20-MAY-2003 (first entry)
XX
DE Human NOV15c cDNA.
XX
KW NOVX; cytostatic; cardiant; antiarteriosclerotic; antiasthmatic; cancer;
KW hypotensive; cardiomyopathy; bronchial asthma; gene therapy; vaccine;
KW human; gene; ss.
XX
OS Homo sapiens.
XX
FN WO200272757-A2.
XX
PD 19-SEP-2002.
XX
PF 08-MAR-2002; 2002WO-US006908.
XX
PR 08-MAR-2001; 2001US-0274101P.
PR 08-MAR-2001; 2001US-0274194P.
PR 08-MAR-2001; 2001US-0274281P.
PR 08-MAR-2001; 2001US-0274322P.
PR 09-MAR-2001; 2001US-0274849P.
PR 12-MAR-2001; 2001US-0275235P.
PR 13-MAR-2001; 2001US-0275578P.
PR 13-MAR-2001; 2001US-0275579P.
PR 13-MAR-2001; 2001US-0275601P.
PR 14-MAR-2001; 2001US-0276000P.
PR 16-MAR-2001; 2001US-0276766P.
PR 19-MAR-2001; 2001US-0276994P.

PR 20-MAR-2001; 2001US-0277239P.
PR 20-MAR-2001; 2001US-0277321P.
PR 20-MAR-2001; 2001US-0277327P.
PR 21-MAR-2001; 2001US-0277791P.
PR 22-MAR-2001; 2001US-0277833P.
PR 23-MAR-2001; 2001US-0278152P.
PR 26-MAR-2001; 2001US-0278894P.
PR 27-MAR-2001; 2001US-0278999P.
PR 27-MAR-2001; 2001US-0279036P.
PR 28-MAR-2001; 2001US-0279344P.
PR 30-MAR-2001; 2001US-0277338P.
PR 30-MAR-2001; 2001US-0279995P.
PR 30-MAR-2001; 2001US-0280233P.
PR 02-APR-2001; 2001US-0280802P.
PR 02-APR-2001; 2001US-0280822P.
PR 02-APR-2001; 2001US-0280900P.
PR 04-APR-2001; 2001US-0281194P.
PR 13-APR-2001; 2001US-0283675P.
PR 30-APR-2001; 2001US-0287424P.
PR 02-MAY-2001; 2001US-0288066P.
PR 03-MAY-2001; 2001US-0288342P.
PR 03-MAY-2001; 2001US-0288528P.
PR 15-MAY-2001; 2001US-0291190P.
PR 16-MAY-2001; 2001US-0291099P.
PR 16-MAY-2001; 2001US-0291240P.
PR 30-MAY-2001; 2001US-0294485P.
PR 31-MAY-2001; 2001US-0294889P.
PR 31-MAY-2001; 2001US-0294899P.
PR 18-JUN-2001; 2001US-0299027P.
PR 19-JUN-2001; 2001US-0299303P.
PR 19-JUN-2001; 2001US-0299310P.
PR 10-JUL-2001; 2001US-0304354P.
PR 31-JUL-2001; 2001US-0309198P.
PR 16-AUG-2001; 2001US-0312903P.
PR 10-SEP-2001; 2001US-0318462P.
PR 12-SEP-2001; 2001US-0318770P.
PR 27-SEP-2001; 2001US-0325430P.
PR 27-SEP-2001; 2001US-0325681P.
PR 18-OCT-2001; 2001US-0330380P.
PR 31-OCT-2001; 2001US-0335301P.
PR 14-NOV-2001; 2001US-0332172P.
PR 14-NOV-2001; 2001US-0332271P.
PR 14-NOV-2001; 2001US-0332272P.
PR 14-NOV-2001; 2001US-0333184P.
PR 21-NOV-2001; 2001US-0333272P.
PR 21-NOV-2001; 2001US-0332094P.
PR 03-DEC-2001; 2001US-0337426P.
PR 03-DEC-2001; 2001US-0338092P.
PR 04-DEC-2001; 2001US-0337185P.
PR 03-JAN-2002; 2002US-0345705P.
PR 07-MAR-2002; 2002US-00092900.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Padigaru M, Spytek KA, Shenoy SG, Taupier RJ, Pena CEA, Li L;
PI Zehrusen BD, Gusev V, Ji W, Gorman L, Miller CE, Kekuda R;
PI Patturajan M, Gangolli E, Vernet CAM, Guo X, Tchernev V;
PI Fernandes ER, Casman SJ, Malyankar UM, Gerlach V, Liu Y, Anderson D;
PI Spaderna SK, Catterton E, Burgess C, Leite M, Zhong H, Alsobrook JP;
PI Lepley DM, Rieger DK;
XX
WPI; 2002-723332/78.
P-PSDB; ABU65069.
XX
DR
DR
XX
PT NOVX polypeptides and polynucleotides, useful for preventing or treating
PT a disorder associated with aberrant NOVX expression or activity e.g.,
PT cancer, hypertension, atherosclerosis, cardiomyopathy or bronchial
PT asthma.
XX
PS Claim 13; Page 139-140; 1103pp; English.
XX
CC This invention describes novel human NOVX polypeptides which have
CC cytostatic, cardiant, antiarteriosclerotic, antiasthmatic and hypotensive

Padigaru M, Spytek KA, Shenoy SG, Taupier RJ, Pena CEA, Li L;
Zehrusen BD, Gusev V, Ji W, Gorman L, Miller CE, Kekuda R;
Patturajan M, Gangolli E, Vernet CAM, Guo X, Tchernev V;
Fernandes ER, Casman SJ, Malyankar UM, Gerlach V, Liu Y, Anderson D;
Spaderna SK, Catterton E, Burgess C, Leite M, Zhong H, Alsobrook JP;
Lepley DM, Rieger DK;
WPI; 2002-723332/78.
P-PSDB; ABU65069.
NOVX polypeptides and polynucleotides, useful for preventing or treating
a disorder associated with aberrant NOVX expression or activity e.g.,
cancer, hypertension, atherosclerosis, cardiomyopathy or bronchial
asthma.
Claim 13; Page 139-140; 1103pp; English.

This invention describes novel human NOVX polypeptides which have
cytostatic, cardiant, antiarteriosclerotic, antiasthmatic and hypotensive

CC activity. Pharmaceutical compositions comprising the NOVX proteins or
CC nucleic acid molecules or NOVX antibodies are useful for preventing or
CC treating a disorder associated with aberrant NOVX expression or activity
CC e.g. cancer, hypertension, atherosclerosis, cardiomyopathy or bronchial
CC asthma. The products of the invention can be used for gene therapy or in
CC a vaccine. ABX97008-ABX97185 are cDNA fragments amplified and isolated by
CC the PCR primers and probes represented in ABX13460-ABX13462 and ABX97186-
CC ABX97593. ABX97008-ABX97185 encode the NOVX proteins described in
CC ABU65041-ABU65218
XX
SQ Sequence 1725 BP; 394 A; 500 C; 523 G; 308 T; 0 U; 0 Other;
Query Match 91.2%; Score 1381.4; DB 6; Length 1725;
Best Local Similarity 92.6%; Pred. No. 0;
Matches 1509; Conservative 0; Mismatches 6; Indels 114; Gaps 1;
QY 1 ATGGAGGGGGTCCAGCTGTCTGCTCCAGAGTCTCTGGGCGAGAGTGTGTAGAACGGGTG 60
DB 52 ATGGAGGGGGTCCAGCTGTCTGCTCCAGAGTCTCTGGGCGAGAGTGTGTAGAACGGGTG 111
QY 61 GCAGCCATCATGTGACTCACTTGGAGGAGGAGATGGTGGCCAGAGCCTACTAGAAAC 120
DB 112 GCAGCCATCATGTGACTCACTTGGAGGAGGAGATGGTGGCCAGAGCCTACTAGAAAC 171
QY 121 GGTGTGAGCCCCCACCACGGGCGAGAGTGCCTCTGTGATCCCTGGCAGTACTTCAAGA 180
DB 172 GGTGTGAGCCCCCACCACGGGCGAGAGTGCCTCTGTGATCCCTGGCAGTACTTCAAGA 231
QY 181 CTGCTCCAGCCCGGCTAGCCTCTCAGCCAGAACTTTCCTACAGAGCGGCCAGCA 240
DB 232 CTGCTCCAGCCCGGCTAGCCTCTCAGCCAGAACTTTCCTACAGAGCGGCCAGCA 291
QY 241 GGAAGCTATCTGAGGCGCAGGCTGGGCTTATGCCACGGGGCTGCCAGCCACATCTCC 300
DB 292 GGAAGCTATCTGAGGCGCAGGCTGGGCTTATGCCACGGGGCTGCCAGCCACATCTCC 351
QY 301 CCCCAGGCTGGCGAGGGCCACCATCGAGTCCCAACAGTGGCCATCTCAGATGAGAG 360
DB 352 CCCCAGGCTGGCGAGGGCCACCATCGAGTCCCAACAGTGGCCATCTCAGATGAGAG 411
QY 361 GACTGGTGCAGTGAACAGTACAGCTCAGAGTGAAGTGGCAGGGTCCCTAGGT 420
DB 412 GACTGGTGCAGTGAACAGTGAACAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 471
QY 421 GTGGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 480
DB 472 GTGGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 531
QY 481 AAAAAGAGTTACTGAAGCAGTATGGCTTTCCACGTGCGCCCTCCCGGAGAGGTCCTCCAG 540
DB 532 AAAAAGAGTTACTGAAGCAGTATGGCTTTCCACGTGCGCCCTCCCGGAGAGGTCCTCCAG 591
QY 541 GCTGCCAGGGGAGCCAGCAGAGCTGCTGCCCTGGAGCGGGTGTACAGGAGATT 600
DB 592 GCTGCCAGGGGAGCCAGCAGAGCTGCTGCCCTGGAGCGGGTGTACAGGAGATT 651
QY 601 GCCATCTGAAGAGTGGACCACTGAATGTGGTCAAACTGATCGAGGTCTCTGGATGAC 660
DB 652 GCCATCTGNAGAGTGGACCACTGAATGTGGTCAAACTGATCGAGGTCTCTGGATGAC 711
QY 661 CCAGCTGAGGACAACTCTATTG- 684
DB 712 CCGGCTGAGGACAACTCTATTG- 771
QY 685 ----- 684
DB 772 TCAACAANTATGCCAAGTCCCACTCCCTGCTTCCCTCTGAGCAGCAGAGAGTGGATCC 831
QY 685 -----GTGTTTGACCTCTCAGAAAGGGGCCGCTCATGGAGTCCCG 726
DB 832 ACGTGGGCTGGGCTCAGTGTGTTGACCTCTGAGAAAGGGGCCGCTCATGGAGTCCCG 891
QY 727 TGTGACAAGCCCTTCTCGGAGGAGCAAGCTGCGCTCTTACCTCGGGGACGTCATCTCGGGC 786

DB 892 TCTGACAAGCCCTTCTCGAGGAGCAAGCTCCTCTACCTCGGGAGCTCATCTCTGGGC 951
QY 787 CTCGAGTACTTCACTGCGCAGAGATCTCCACAGGAGCATCAAGCCATCCAACTGCTC 846
DB 952 CTCGAGTACTTCACTGCGCAGAGATCTCCACAGGAGCATCAAGCCATCCAACTGCTC 1011
QY 847 CTGGGGGATGATGGGCACGTGAAGATCGCGACTTTGGCGTCAGCAACACAGTTTGAGGGG 906
DB 1012 CTGGGGGATGATGGGCACGTGAAGATCGCGACTTTGGCGTCAGCAACACAGTTTGAGGGG 1071
QY 907 AACAGCGCTCAGCTGTCCAGACGCGGGGAAACCCAGCATTTTCATGCCCCCGAGGCAATT 966
DB 1072 AACAGCGCTCAGCTGTCCAGACGCGGGGAAACCCAGCATTTTCATGCCCCCGAGGCAATT 1131
QY 967 TCTGATTCGGGCGCAGAGCTTCAGTGGGAAGGCTTCAGTGGATGTATGGCCACTGCGCTCAGC 1026
DB 1132 TCTGATTCGGGCGCAGAGCTTCAGTGGGAAGGCTTCAGTGGATGTATGGCCACTGCGCTCAGC 1191
QY 1027 TTGTACTGCTTTTGTCTATGGGAAGTSCCAATTCATCGACGATTTTCATCTGGCCCTCCAC 1086
DB 1192 TTGTACTGCTTTTGTCTATGGGAAGTSCCAATTCATCGACGATTTTCATCTGGCCCTCCAC 1251
QY 1087 AGAAGATCAAGATGAGCCCGTGGTTCCTGAGGAGCCAGAAATCAGCAGGAGCTC 1146
DB 1252 AGAAGATCAAGATGAGCCCGTGGTTCCTGAGGAGCCAGAAATCAGCAGGAGCTC 1311
QY 1147 AAGGACCTCATCTGAAGATTTAGACAAGATTCAGCAGAGATTCAGGAGTCCAGAC 1206
DB 1312 AAGGACCTCATCTGAAGATTTAGACAAGATTCAGCAGAGATTCAGGAGTCCAGAC 1371
QY 1207 ATCAAGTTGACCCCTTGGGTGACCAAGAACGGGGAGAGCCCTTTCCTTCGAGGAGGAG 1266
DB 1372 ATCAAGTTGACCCCTTGGGTGACCAAGAACGGGGAGAGCCCTTTCCTTCGAGGAGGAG 1431
QY 1267 CACTGAGCGTGGTGGAGGTGACAGAGGGGAGGTTAAGACTCAGTCAAGCTCATCCCTC 1326
DB 1432 CACTGAGCGTGGTGGAGGTGACAGAGGGGAGGTTAAGACTCAGTCAAGCTCATCCCTC 1491
QY 1327 AGCTGACACACGCTGATCTGGTGAAGTCCATGCTCAGGAAGCGTTTCCTTTGGGAACCCG 1386
DB 1492 AGCTGACACACGCTGATCTGGTGAAGTCCATGCTCAGGAAGCGTTTCCTTTGGGAACCCG 1551
QY 1387 TTTGAGCCCCCAGCACGAGGAGGAGCGATCATCTGTCTCCAGAAACCTTACTGGTG 1446
DB 1552 TTTGAGCCCCCAGCACGAGGAGGAGCGATCATCTGTCTCCAGAAACCTTACTGGTG 1611
QY 1447 AAGAGAGGTTGGTGAAGGGGCAAGCCAGAGCTCCCGGCGTCCAGGAGAGCGAG 1506
DB 1612 AAGAGAGGTTGGTGAAGGGGCAAGCCAGAGCTCCCGGCGTCCAGGAGAGCGAG 1671
QY 1507 GCTGCATCC 1515
DB 1672 GCTGCATCC 1680
RESULT 10
ADB53308
ID ADB53308 standard; DNA; 3411 BP.
XX ADB53308;
XX
XX
XX 04-DEC-2003 (first entry)
DE Primary rat hepatocyte toxicity modelling related gene SEQ ID NO:3850.
XX toxic effect; gene expression profile; hepatotoxicity; diagnostic marker;
KW toxicity marker; toxicity progression; drug screening;
KW primary rat hepatocyte toxicity modelling; gene; ds.
XX
OS Rattus norvegicus.
XX
FN W02003065993-A2.

XX 14-AUG-2003.
XX
XX
XX 04-FEB-2003; 2003WO-US003482.
XX
XX 04-FEB-2002; 2002US-0353171P.
PR 13-MAR-2002; 2002US-0363534P.
PR 08-APR-2002; 2002US-0370248P.
PR 10-APR-2002; 2002US-0371134P.
PR 10-APR-2002; 2002US-0371135P.
PR 10-APR-2002; 2002US-0371150P.
PR 11-APR-2002; 2002US-0371413P.
PR 19-APR-2002; 2002US-0373601P.
PR 19-APR-2002; 2002US-0373602P.
PR 22-APR-2002; 2002US-0374139P.
PR 08-MAY-2002; 2002US-0378370P.
PR 09-MAY-2002; 2002US-0378652P.
PR 09-MAY-2002; 2002US-0378653P.
PR 09-MAY-2002; 2002US-0378665P.
PR 09-JUL-2002; 2002US-0394230P.
PR 09-JUL-2002; 2002US-0394253P.
PR 04-SEP-2002; 2002US-0407688P.
PR 28-JAN-2003; 2003US-0442900P.
XX
XX (GENE-) GENE LOGIC INC.
XX
XX Mendrick D, Porter M, Johnson K, Higgs B, Castle A, Orr M;
PI Elashoff M;
XX
XX WPI; 2003-731472/69.
XX
XX Determining if a compound induces a toxic effect on a tissue or cell, for
PT identifying hepatotoxic compounds, comprises comparing a gene expression
PT profile of a tissue or cell sample to a database of tox mean and non-Tox
PT mean values.
XX
XX Claim 44; SEQ ID NO 3850; 874pp; English.
XX
XX The present invention describes a method for determining whether a
CC compound induces a toxic effect on a tissue or cell. The method comprises
CC preparing a gene expression profile of a tissue or cell sample exposed to
CC the compound, and comparing the gene expression profile to a database
CC comprising data or information on the tox mean and non-Tox mean value.
CC The method is useful for predicting or identifying at least one toxic
CC effect, particularly hepatotoxicity, of a test or unknown compound. The
CC genes listed in the specification are useful as diagnostic or toxicity
CC markers for the prediction or identification of the physiological state
CC of tissue or cell sample that has been exposed to a compound, or to
CC identify or predict the toxic effects of a compound or an agent. These
CC may also be used as markers for monitoring toxicity progression or for
CC drug screening. The present sequence represents a primary rat hepatocyte
CC toxicity modelling related gene sequence from the present invention.
XX
XX Sequence 3411 BP; 823 A; 875 C; 968 G; 745 T; 0 U; 0 Other;
XX
XX Query Match 81.6%; Score 1236.6; DB 9; Length 3411;
XX Best Local Similarity 88.5%; Pred. No. 3.8e-298;
XX Matches 1341; Conservative 0; Mismatches 174; Indels 0; Gaps 0;
XX
XX 1 ATGAGGGGGGTTCAGCTGTCTGTCAGAGATCCTCGGGCAGAGCTGGTAGAACGGGTG 60
DB 107 ATGAGGGCGATCCAGCGCTGTGTCACAGACCTCGAGCGGAACCTGGTAGCGGGTG 166
XX
XX 61 GCACCATCATGTGACTCACTTGGAGGAGCAGATGTGCGCCAGAGCCCTACTAGAAC 120
DB 167 GCACCATCATGTGGCCCACTAGAGAGCAGAGGGGCCAGAGCCCTGCGAGCAAT 226
XX
XX 121 GGTGTGAGACCCCAACACGGGGCAGAGCTGCCTCTGTGATCCCTGGCAGTACTTCAAGA 180
DB 227 GGTGTGAGACCCCAACACGGGGCAGAGCTGCCTCTGTGATCCCTGGCAGTACTTCAAGA 286
XX
XX 181 CTGCTCCACCGCCGCTACCTCTGACCGAGGAGCTTCCCTTACAGGAGCGGCCAGCA 240
DB

Db 287 CCCACACAGTGGTCCCGAGCCTCTCTGTAGAAAGTTCTCCTCGAGGAAGACCGAG 346
QY 241 GGAAGCTATCTGGAGGCGCAGGCTTGGGCCCTTATGCGACGGGGCCTGCCAGCACATCTCC 300
Db 347 GGAAGCTGTCTAGAGGCTCAGGTTGGGCCCTTACTCTACAGGACCTGCCAGTACATGTCT 406
QY 301 CCCCGGGCTGGCGAGGCGCCACCATCGAGTCCCAACCATCGAGTGGCCATCTCAGATGAGAG 360
Db 407 CCTCGGGCTTGGCGGAGAGCCCAACCATCGAGTCCCAACCATCGAGTGGCCATCTCAGACACAGAG 466
QY 361 GACTCGGTGAGCTCAAGTACAGTACAGTGGAGTGGAGTGGCAAGGTGCCACGCT 420
Db 467 GACTGTGTCAACTGAACAGTACAGTGGAGTGGAGTGGCAAGGTGCCATGTATGT 526
QY 421 GTGTGAGGCTGGCCCTACAAAGAAAGTGAAGACAGACACTATGCAATGAAAGTCTCTTCC 480
Db 527 GTGTGAGGCTGGCCCTACAAAGAAAGGGAAGACAGACACTATGCAATGAAAGTCTCTTCC 586
QY 481 AAAAAGAAAGTACTGAAGCAGTATGGCTTCCAGTGGCCCTCCCGAGAGGGTCCGAG 540
Db 587 AAAAAGAAAGTACTGAAGCAGTATGGCTTCCAGTGGCCCTCCCGAGAGGGTCCCAA 646
QY 541 GCTGCCAGGAGGAGCAGCCAGCAGCTGTGCGCCCTGGAGCGGGTGTACCCAGAGATT 600
Db 647 GCTCTCAGGAGGAGCCAGCCAAACAGCTGTGCCCCCTGGAGCGTGTATACCAGGAGATT 706
QY 601 GCCATCTGAAGAAAGCTGGACCCACGTAATGTGTGTTGTTCAAACTGATCGAGTCTCTGGATGAC 660
Db 707 GCCATTTTAAAGAAACTGGACCCACGTAATGTGTGTTGTTCAAGTTGATCGAGTCTCTGGATGAT 766
QY 661 CCAGTGGAGCAACCTCTATTTGGTGTGTTGACCTCTCGAGAAAGGGGCCCTCATGAA 720
Db 767 CCTGTGAAGCAATCTCTATTTGGTGTGTTGACCTCTCGAGAAAGGGACCACTCATGGA 826
QY 721 GTGCCCTGTGAAGCCCTTCTCGAGGAGCAAGCTCGCCCTTACTCTCGGAGCTCATC 780
Db 827 GTGCCCTCGACAAAGCCCTTCCAGAGAGCAAGCTCGCCCTTACTCTCGGAGCAATCATC 886
QY 781 CTGGCCCTCGAGTACTTGTCACTGCGAGAAAGATCGTCCACAGGACATCAAGCCATCCAAAC 840
Db 887 CTGGCCCTCGAGTACTTGTCACTGCGAGAAAGATTTGTCACAGGACATCAAGCCGTTCAAT 946
QY 841 CTGCTCTCTGGGGGATGATGGCAGTGAAGATCCCGACTTGGCGTTCAGCAACCAAGTTT 900
Db 947 CTGCTCTCTGGGAGATGGGCAGTGAAGATGCCGACTTGGTGTTCAGCAACCAAGTTT 1006
QY 901 GAGGGGAACGAGCTCAGCTGTCCAGCAGCGCGGGAACCCAGCAATTCATGGCCCCGAG 960
Db 1007 GAGGGGAATGATGCTCAGCTGTCCAGTACGCGAGGACCCAGCAATTCATGGCCCCGAG 1066
QY 961 GCCATTTCTGATTCGGGCCAGAGCTTCAGTGGGAAGGCTTGGATGTATGGCCACCTGGC 1020
Db 1067 GCCATCTCTGACACCGGCCAGAGCTTCAGTGGGAAGGCTTGGATGTATGGCCACCTGGG 1126
QY 1021 GTACAGTGTACTGCTTGTCTATGGGAAGTGGCCATTCATCGACGATTTTCATCTGCGC 1080
Db 1127 GTTACATGTATGTTTGTCTACGGAAGTGGCCATTCATGTAGTATCATCTGCGC 1186
QY 1081 CTCACAGGAAGATCAAGAAATGAGCCCGTGGTGTTCCTCGAGGAGCCAGAAATCAGCGAG 1140
Db 1187 CTGCACAAAGATCAAGAAATGAGCCCGTGGTGTTCCTCGAGGAGCCGAGGTTCAGCGAG 1246
QY 1141 GAGCTCAGGACCTGATCTCTGAGATGTGTAGACAGAAATCCGAGAGAGAAATTTGGGGTG 1200
Db 1247 GAATCTAAAGACCTGATCTCTGAAGATGTGTAGCAAGAAATCTTGAACAAGAAATTTGGGGTG 1306
QY 1201 CCAGACATCAAGTGTGACCTTGGGTGACCAAGAAACGGGAGGAGCCCTTCTCTTCGGAG 1260
Db 1307 TCTGATATCAAGTTACACCTTGGGTGACCAAGCAATCGAGAGGAGCCCTTCTCTTCAGAG 1366
QY 1261 GAGAGACATGCGCGTGTGTGGAGGTGACAGAGGGGAGGTTAAGAACTCAGTCAGGCTC 1320
Db 1367 GAGAACACTGCAGTGTGTGTAGAGGTGACTGAGGAGGAGGTGAAGAACTCAGTCAAGCTC 1426

QY 1321 ATCCCAGCTGGACACGGTGATCCCTGGTGAAGTCCATGCTGAGGAAGCTTCTCTTTGGG 1380
Db |||||
QY 1427 ATCCCAGCTGGACCACTGTGATCCTGGTCAAGTCTATGCTGAGAAAGCGTTTCCTTTGGA 1486
Db |||||
QY 1381 AACCCCTTTGAGCCCCAGGACGCGAGGGAAGACGATCCATGCTGCTCCAGGAAACCTTA 1440
Db |||||
QY 1487 AACCCATTTGAGCCCCAAGCAGCAGGGAAGAAAGATCTATGCTGCGGCCAGGAAACTTA 1546
Db |||||
QY 1441 CTGGTGAAGCAAGGGTTTGTGAAGGGGGCAAGGCCAGAGCTCCCCGGCGTCCAGGAA 1500
Db |||||
QY 1547 CTGTTGAAGAAAGAGATGTGAGAAAGGGGGCAAAAGCCAGAGCTTCCCGAGTCCAGGAA 1606
Db |||||
QY 1501 GACGAGGCTGCATCC 1515
Db |||||
QY 1607 GATGAGGCTGCATCC 1621
Db |||||

RESULT 11

AAZ29223

ID AAZ29223 standard; cDNA; 2545 BP.

XX

AC AAZ29223;

XX

DT 28-FEB-2000 (first entry)

XX

XX Human cell signalling protein-2 encoding cDNA.

XX

KW Cell signalling protein-2; CSIGP-2; cell proliferation; arteriosclerosis;
KW inflammatory disorder; cirrhosis; cancer; hepatitis; AIDS;
KW Addison's disease; multiple sclerosis; ss.

XX

OS *Homo sapiens*.

XX

FH Key Location/Qualifiers

FT CDS 241..1860

FT /*tag= a

FT /product= "Cell Signalling Protein-2"

XX

XX WO9558558-A2.

XX

PD 18-NOV-1999.

XX

XX 13-MAY-1999; 99WO-US010567.

XX

XX 13-MAY-1998; 98US-0085343P.

PR

XX 26-AUG-1998; 98US-0098010P.

XX

PA (INCY-) INCYTE PHARM INC.

XX

PI Bandman O, Hillman JL, Lal P, Yue H, Tang YT, Patterson C;

PI Baughn MR, Yang J;

XX

XX WPI; 2000-086432/07.

DR

XX P-PSDB; AAY44239.

XX

XX Human cell signaling proteins useful for, e.g. diagnosing cell

XX

XX proliferative and inflammatory disorders.

XX

XX Claim 9; Page 80-81; 90pp; English.

XX

CC The present sequence is a cDNA obtained from Incyte clone 640521 of
CC BRSTN03 library. It encodes cell signalling protein-2 (CSIGP-2). It is
CC expressed in reproductive, nervous and developmental tissues. Fragments
CC of CSIGP encoding nucleic acid can be used as hybridisation probe for
CC detecting CSIGP related sequences or allelic variants. Recombinant CSIGP
CC can be produced in host cells by transforming them with genetically
CC engineered vectors. Agonists or antagonists can be used in the treatment
CC of cell proliferative and inflammatory disorders associated with
CC decreased or increased CSIGP expression. CSIGP is used in the diagnosis,
CC prevention and treatment of cell proliferative disorders like
CC arteriosclerosis, cirrhosis, cancer, hepatitis and inflammatory disorders
CC like AIDS, Addison's disease, multiple sclerosis, etc

Db 1578 CACGAGGCATGGGCGGAGCGCTTGCCTGGAGGATGAGAACTGCACCGTGGTCGAAGT 1637
QY 1287 GACAGAGGGAGGTTAAGAACTCAGTCAGGCTATCCCGAGCTGACCAACGAGTATCT 1346
Db 1638 GACTGAAGAGGAGGTCGAGAACTCAGTCAAAACACATTCACGCTTGGCAACGCTGATCCT 1697
QY 1347 GGTGAAGTCCATGCTGAGGAAGGTTCTTTTGGGAACCCGTTTGAGCCCCAGGCGGAG 1406
Db 1698 GGTGAAGACCATGATACGTAAAGCTCTTTTGGGAACCCATTCGAGGGC---AGCGGGC 1754
QY 1407 GGAAGAGCATCATGTCTGCTCCAGGAAACCTACTGTGGAAGAGGTTTGGTGAAGG 1466
Db 1755 CGAGGAACGCTCACTGTACAGCGCTCGAAACTTGCTCAGGAAGGAGGCGAGCAAGAA 1814
QY 1467 GGGCAGAGCCAGAGCTCCCGGGGCTCCAGGAAGACGAGG 1507
Db 1815 CCTCAGGGCACCGACCCCGCCCGGTGGGGGAGGAGG 1855

RESULT 12
ADB75227
ID ADB75227 standard; cDNA; 4942 BP.

XX ADB75227;

AC 04-DEC-2003 (first entry)

DT Prostate cancer marker cDNA.

DE Prostate; cancer; cytostatic; gene therapy; marker; ss.

DE Homo sapiens.

OS WO2003009814-A2.

PN 06-FEB-2003.

XX 25-JUL-2002; 2002WO-US023913.

XX 25-JUL-2001; 2001US-0307982P.

PR 22-AUG-2001; 2001US-0314356P.

PR 25-SEP-2001; 2001US-0325020P.

PR 12-DEC-2001; 2001US-0341746P.

PR 05-MAR-2002; 2002US-0362158P.

XX (MILL-) MILLENNIUM PHARM INC.

XX Schlegel R, Monahan JE, Endege WO, Gannavarapu M, Gorbacheva B;

PI Hoersh S, Kamatkar S, Wonsey AM, Glatt K, Zhao X, Anderson D;

XX WPI; 2003-248033/24.

XX New nucleic acid molecule, useful for diagnosing or treating prostate

PT cancer.

XX Claim 1; SEQ ID NO 51; 99pp; English.

XX The invention relates to newly discovered cancer markers associated with

CC the cancerous state of prostate cells. Also disclosed is a method of

CC assessing whether a patient is afflicted with prostate cancer. The method

CC of the invention involves assessing whether a patient is afflicted with

CC prostate cancer by comparing the level of expression of a marker in a

CC patient sample and the normal level of expression of the marker in a

CC control non-prostate cancer marker, where a significant increase in the

CC level of expression of the marker in the patient sample and the normal

CC Nucleic acids of the invention are useful for diagnosing or treating

CC prostate cancer, and may be useful in gene therapy. Sequences given in

CC ADB75177-ADB75631 represent marker cDNA and proteins. Note: The sequence

CC data for this patent did not form part of the printed specification, but

CC was obtained in electronic format directly from WIPO at

CC ftp.wipo.int/pub/published_pct_sequences.

XX

SQ Sequence 4942 BP; 1169 A; 1294 C; 1356 G; 1122 T; 0 U; 1 Other;
Query Match 43.0%; Score 651.2; DB 9; Length 4942;
Best Local Similarity 71.3%; Pred. No. 5e-152;
Matches 873; Conservative 0; Mismatches 348; Indels 3; Gaps 1;
QY 267 GCCTTATGCCACGGGCGCTGCAGCCACATCTCCCGCGGCGCTGGCGGAGGCCACCAT 326
Db 536 GCCCTACTCTACCCGTCAGCTCCCGCAGTCTCGCTCGGCTGCCCGCGGCGGACAGT 595
QY 327 CGAGTCCCAACACGCTGGCCATCTCAGATCAGAGGACTGCGTGCAGCTGAACAGTACAA 386
Db 596 GAGTCTCACCAGCTCTCCATCACCAGTATGACAGACTGTGTGAGCTGAATCAGTATAC 655
QY 387 GCTCAGAGTGAATGGCAAGGTCCTTACGCTGTGTGAGGCTGGCCCTACAAAGAAAG 446
Db 656 CCTGAAGGATGAAATTGGAAAGGGCTCTATGCTGCTCAAGTTGGCTTACAATGAAA 715
QY 447 TGAAGACACACACTATGCAATGAAAGTCTTTTCCAAAAGAGTTACTGAAAGCATATGG 506
Db 716 TGACAATCTACTATGCAATGAAAGTCTGTGCAAAAAGAGTGTATCCGCGAGCCCGG 775
QY 507 CTTTCCAGCTCGCCCTCCCGCAGAGGCTCCAGGCTGCCAGGAGGACCAAGCA 566
Db 776 CTTTCCAGCTCGCCCTCCAGCAGGACCCCGCAGCTCCTGGAGGCTGCATCCAGCC 835
QY 567 GCTGCTGCCCTGGAGCGGCTGTACAGAGATGCCATCTCTGAAAGAGCTGGACCACT 626
Db 836 CAGGGCCCCCATTTGAGCAGGTTGTACAGGAAATGCCATCTCTCAAGAAAGTGGACACC 895
QY 627 GAATGTGTCAAACCTGATCGAGGCTCTCGATGACCCAGCTGAGGACCACTTATTTGT 686
Db 896 CAATGTGTGAAGCTGTGTGAGGCTCTCGATGACCCCAATGAGGACCATCTGTACATGT 955
QY 687 GTTTGACCTCTGAGAAAGGGCGCTCATGGAAGTGCCTGTGCAAGACCCCTTCTCGA 746
Db 956 GTTGAACCTGGTCAACCAAGGGCCGCTGATGGAAGTGCCTCAACCACTCTCTGA 1015
QY 747 GGAGCAAGCTCGCTCTACCTTGGGAGCGTCACTCTGGGCTCGAGTACTTGCATGGCA 806
Db 1016 AGACCAAGCGGCTTTCTACTTCCAGGATCTGATCAAGAGCATCGAGTACTTACACTCA 1075
QY 807 GAAGATCGTCACAGGACATCAAGCCATCAACCTGCTCTCTGGGGGATGATGGCAGCT 866
Db 1076 GAAGATCATCACCGTGACATCAAACTTCAACCTCTCTGGTCGAGAGATGGGACAT 1135
QY 867 GAAGATCGCCGACTTTTGGCTGACCAACAGTTTGGAGGGAACGACGCTCAGCTCCAG 926
Db 1136 CAAGATCGCTGACTTTGGTGTGAGCAATGAATTCAGAGGCGAGTGACGCGCTCTCTCAA 1195
QY 927 CACGGGGGAACCCAGCATTTGCGGCCCGGAGGCCATTTCTGATTCGGGCGCAGGCTT 986
Db 1196 CACGCTGGGCGACGCCCGCTTCTGAGCAACCGAGTCTCTCTGAGAGCCCGCAAGATCTT 1255
QY 987 CAGTGGGAAGGCTTTGGATGTATGGCCACTGGCGTCACTTGTACTCTTTGCTATGG 1046
Db 1256 CTCTGGGAGGCTTTGATGTTTGGCCATGGGTGTGACATATACCTGCTTTGCTTTGG 1315
QY 1047 GAATGCCCATTCATCGAAGATTTTCACTCTGGCCCTCCACAGGAAGATCAAGAATGAGCC 1106
Db 1316 CCAATGGCCATTCATGACGAGCGGATCATGTGTTTACACAGTAAGATCAAGAGTCAGGC 1375
QY 1107 GGTGCTGTTTCTCAGGAGCCAGAAATCAGCGAGGAGCTCAAGACCTGATCTCTGAAGAT 1166
Db 1376 CCTGGAATTTCCAGCAGCCCGGACATAGCTGAGGACTTGAAGGACCTGATCAACCCG 1435
QY 1167 GTTAGACAAGATCCCGAGAGAGATTTGGGGTCCCGAGATCAAGTTGCAACCTTTGGGT 1226
Db 1436 GCTGCAACAAGAACCCCGAGTGCAGGATCGTGTGCTCCGGAATCAAGCTGACCCCTGGGT 1495
QY 1227 GACCAAGAAAGGGGAGGAGCCCTTCTTCCGAGGAGGAGCACTGACGCTGGTGGAGGT 1286
Db 1496 CACGAGGCATGGGGCGGAGCGCTTGGCGTCGAGGATGAGAACTGCACGCTGGTCAAGT 1555

QY	1287	GACAGAGGGGAGGTTAAGAACTCAGTCAGGCTCATCCGAGCTGGACCAACGGTGATCCT	1346
Db	1556	GACTGAAGAGGAGGTCGAGAACTCAGTCAAAACATTCGCCAGCTTGCCAAACCGTGATCCT	1615
QY	1347	GGTGAAGTCCATGCTGAGGAAGGTTCCCTTTGGGAACCCGTTTGAGCCCCCAGGCACGGAG	1406
Db	1616	GGTGAAGACCATGATACGTAACGCTCCTTTGGGAACCCCATTCGAGGCG---AGCCGGCG	1672
QY	1407	GGAGAGCGGATCCATGTCCTCTCCAGGAAACCTACTGGTGAAAGAGGGTTTGGTGAAGG	1466
Db	1673	GGAGGAACGCTACTGTCAGCGCTGGAAACTTGCCTCACCACCAACCAACGAGGAATG	1732
QY	1467	GGGCAAGAGCCAGAGCTCCGCCG 1490	
Db	1733	TGAGTCCCTGCTGAGCTCAAGG 1756	
RESULT 13			
ABK92170			
ID	ABK92170 standard; DNA; 1804 BP.		
XX			
AC	ABK92170;		
XX			
DT	15-AUG-2002 (first entry)		
XX			
DE	Prostate cancer-associated DNA sequence #56.		
XX			
KW	Prostate cancer; prostate tumour tissue; human; mammal; cytostatic;		
KW	gene therapy; Gene; ds.		
XX			
OS	Mammalia.		
XX			
PN	WO200230268-A2.		
XX			
PD	18-APR-2002.		
XX			
PF	12-OCT-2001; 2001WO-US032045.		
XX			
PR	13-OCT-2000; 2000US-00687576.		
PR	08-DEC-2000; 2000US-00733288.		
PR	08-DEC-2000; 2000US-00733742.		
PR	24-JAN-2001; 2001US-0263957P.		
PR	16-MAR-2001; 2001US-0276791P.		
PR	16-MAR-2001; 2001US-0276888P.		
PR	06-APR-2001; 2001US-0281922P.		
PR	24-APR-2001; 2001US-0286214P.		
PR	30-APR-2001; 2001US-00847046.		
PR	04-MAY-2001; 2001US-0288589P.		
XX			
PA	(EOSB-) EOS BIOTECHNOLOGY INC.		
XX			
PI	Gish KC, Mack DH, Wilson KE, Afar D, Hevezi P;		
XX			
XX	WPI; 2002-471335/50.		
DR	P-PSDB; ABG61855.		
XX			
PT	Detecting a prostate cancer-associated transcript in a cell in a patient;		
PT	useful for diagnosing prostate cancer (PC) or screening modulators of PC;		
PT	by determining if prostate cancer-associated genes are expressed in a		
PT	prostate tissue.		
XX			
PS	Claim 22; Page 342; 436pp; English.		
XX			
CC	The present invention relates to methods of detecting a prostate cancer-		
CC	associated transcript in a cell from a patient. The method comprises		
CC	contacting a biological sample from the patient with prostate cancer-		
CC	associated polynucleotides (designated PC genes) that selectively		
CC	hybridise to a sequence that is at least 80% identical to them. The		
CC	prostate cancer-associated polynucleotide sequences are differentially		
CC	expressed in prostate tumour tissue or in prostate cancer and are derived		
CC	from the tissues of various organisms such as humans or other mammals		
CC	(e.g. mice, sheep and dogs). The methods of the invention are useful for		

CC	diagnosing and treating prostate cancer in mammals. The prostate cancer-
CC	associated genes are useful for diagnosing or treating prostate cancer,
CC	as well as for identifying modulators of prostate cancer or agents that
CC	inhibit prostate cancer. The nucleic acid sequences are particularly
CC	useful in gene therapy, as a vaccine or in antisense applications.
CC	ABK92115-ABK92263 represent prostate cancer-associated polynucleotide
CC	sequences
XX	
SEQ	Sequence 1804 BP; 445 A; 519 C; 453 G; 387 T; 0 U; 0 Other;
	Query Match 43.0%; Score 651; DB 6; Length 1804;
	Best Local Similarity 71.5%; Pred. No. 3.9e-152;
	Matches 871; Conservative 0; Mismatches 345; Indels 3; Gaps 1;
QY	267 GCCTTATGCCACGGGCGCTGCCAGCCACATCTCCCCCGGCGCTGGCGGAGGCCACCAT 326
DB	30 GCCTACTACCCGTGAGTCCCCGCGAGTCTCGCCTCGCTGCGCGCGCGCGACAGT 89
QY	327 CGAGTCCGACCGTGGCCATCTCAGATCAGAGGAGTGGTGAGCTGAACAGTACAA 386
DB	90 GGAAGTCTCACCGAGTCTCCATCAGGATATGACAGGACTGTGTCAGCTGAATCAGTATAC 149
QY	387 GCTGCAGAGTGAGATTGGCAAGGCGTCTACGCTGTGGTGCAGGCTGCCTACACGAAG 446
DB	150 CCTGAAGGATGAATTTGGAAGGCTCCTATGGTGTGCTCAAGTTGGCTTACATGAAA 209
QY	447 TGAAGACAGACACTATGCAATGAAAGTCTTTCCAAAAGAAGTTACTGAAGCAGTATGG 506
DB	210 TGACATACCTACTATGCAATGAAGTGTGTCCAAAAGAAGCTGATCGGCAGGCGCG 269
QY	507 CTTTCCACGTGCGCCCTCCCGGAGAGGTCCTCAGGCTGCCAGGGAGGACGAGCCAGCA 566
DB	270 CTTTCCACGTGCGCCCTCCACCCGAGGACACCGGCCAGCTCCTGGAGGCTGCATCCAGCC 329
QY	567 GCTGCTGCCCTGGAGCGGTGTACAGGAGATTGCCATCTGAAGAAGCTGGACACGT 626
DB	330 CAGGGGCCCCATTGAGCAGGTGTACAGGAAATGGCCATCCTAAGAAGCTGGACCAACC 389
QY	627 GAATGTGGTCAAACTGATCGAGGTCTCTGGATGACCCAGCTGAGGACAACTCTATTGGT 686
DB	390 CAATGTGTGAAGCTGTGTGAGGTCTTGATGACCCCAATCAGGACCATCTGTACATGGT 449
QY	687 GTTTGACCTCCTGAGAAAGGGCCGCTCATGGAAGTGCCTGTGCAAGGCCCTTCTCGGA 746
DB	450 GTTCGAACCTGGTCAACCAAGGGCCGCTGATGGAAGTGCCACCCCTCAAAACCACTCTCTGA 509
QY	747 GGAGCAAGCTCGCTCTACCTGCGGAGCTCATCTGGGCTCGAGTACTTGCACTGCCA 806
DB	510 AGACCAGGCCCGTTTCTACTTCCAGGATCTGATCAAAAGGCATCAGTACTTACACTACCA 569
QY	807 GAAGATCGTCCACAGGACATCAAGCCATCCAACTGCTCTGGGGATGATGGGCACGT 866
DB	570 GAAGATCATCCACGTTGACATCAACCTTCCAACTCTCTGGTCGAGAAAGTGGGCACAT 629
QY	867 GAAGATCCGCGACTTTGGGCTCAGCAACCAAGTTTGAGGGGAAACGACCTCAGCTGTCAAG 926
DB	630 CAAGATCGCTGACTTTGGTGTGAGCAATGAATTCAGGGCAGTGACGCGTCTCTCTCGAA 689
QY	927 CAGGGCGGGAACCCACAGTTCATGGCCCCCGAGGCCATTCTGATTCCGGCCGAGCTT 986
DB	690 CACGTTGGGACGCGCCGCTTCATGGCACCCGAGTCGCTCTCTGAGACCCGCAAGATCTT 749
QY	987 CAGTGGGAAGGCCCTTGAGATGTATGGGCCACTGGCGTTCACCTGTACTGCTTTGTCTATGG 1046
DB	750 CTCTGGGAAGGCCCTTGAGATGTTTGGGCCCATGGGTGACACTACTCTGCTTTGCTTTGG 809
QY	1047 GAAGTGCCCAATTCATCGACGATTTTCATCTCGGCCCTCCACAGGAAGATCAAGAAATGAGCC 1106
DB	810 CCAGTGGCCCAATTCATGACGAGCGGATCATGTGTTTACAGTAAGATCAAGAGTCAAGC 869
QY	1107 CGTGGTGTTCCTGAGGAGCCAGAAATCAGCGGAGAGCTCAGGACCTGATCCTGAAGAT 1166
DB	870 CCTTGGAAATTTCCAGACAGCCGACATAGCTGAGGACTTGAAGGACCTGATCACCCGAT 929

QY 1167 GTTAGCAGAAATCCGACAGAGAAATGGGTCAGACATCAAGTTGCACCTTGGGT 1226
Db 930 GTTGACAGAAACCCGAGTCAGGATCGTGTGCGGAAATCAAGCTGCACCCCTGGT 989
QY 1227 GACCAAGAACGGGAGGAGCCCTTCCCTTCGGAGGAGGAGCACTGCAGCGTGTGGAGT 1286
Db 990 CACGAGGCATGGGCGGAGCCGTTGCCGTCGAGGATGAGAACTGCACGCTGTGAGT 1049
QY 1287 GACAGAGGGGAGGTTAGAACTCAGTCAGGCTCATGCCAGTCATGCCAGTCGACACGTCATCCT 1346
Db 1050 GACTGAAGAGGAGGTCGAGAACTCAGTCAAAACACATTCACAGCTTGGCAACCGTGATCCT 1109
QY 1347 GGTGAAGTCCATGCTGAGAGAGCGTTCCTTTGGGAACCGTTTGGAGCCCGGAGG 1406
Db 1110 GGTGAAGACCATGATACGTAAACGCTCTCTTTGGGAACCGTTTGGAGGCG --AGCCGCG 1166
QY 1407 GGAAGAGCGATCCATGCTGCTCCAGGAAACCTACTGTGAAAGAGGTTTGGTGAAG 1466
Db 1167 GGAGGAACGCTCACTGTGAGGCGCTGGAACTTGTCTACCAAAAACCAACGAGGAATG 1226
QY 1467 GGGCAGAGCCAGAGCTC 1485
Db 1227 TGAGTCCCTGTGAGCTC 1245

RESULT 14
ADB75229
ID ADB75229 standard; cDNA; 1804 BP.
AC ADB75229;
XX
DT 04-DEC-2003 (first entry)
XX
DE Prostate cancer marker cDNA.
XX
KW Prostate; cancer; cytostatic; gene therapy; marker; ss.
OS Homo sapiens.
XX
XX WO2003009814-A2.
XX
PD 06-FEB-2003.
XX
PF 25-JUL-2002; 2002WO-US023913.
XX
PR 25-JUL-2001; 2001US-0307982P.
PR 22-AUG-2001; 2001US-0314356P.
PR 25-SEP-2001; 2001US-0325020P.
PR 12-DEC-2001; 2001US-0341746P.
PR 05-MAR-2002; 2002US-0362159P.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX

PI Schlegel R, Monahan JE, Endege WO, Gannavarapu M, Gorbacheva B;
PI Hoersh S, Kamatkar S, Woney AM, Glatt K, Zhao X, Anderson D;
XX
XX WPI; 2003-248033/24.
XX

XX New nucleic acid molecule, useful for diagnosing or treating prostate
XX cancer.
PT

XX Disclosure; SEQ ID NO 53; 99pp; English.
PS

XX The invention relates to newly discovered cancer markers associated with
CC the cancerous state of prostate cells. Also disclosed is a method of
CC assessing whether a patient is afflicted with prostate cancer. The method
CC of the invention involves assessing whether a patient is afflicted with
CC prostate cancer by comparing the level of expression of a marker in a
CC patient sample and the normal level of expression of the marker in a
CC control non-prostate cancer sample, where a significant increase in the
CC level of expression of the marker in the patient sample and the normal
CC level indicates that the patient is afflicted with prostate cancer.

CC Nucleic acids of the invention are useful for diagnosing or treating
CC prostate cancer, and may be useful in gene therapy. Sequences given in
CC ADB75177-ADB75631 represent marker cDNA and proteins. Note: The sequence
CC data for this patent did not form part of the printed specification, but
CC was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX

SQ Sequence 1804 BP; 445 A; 519 C; 453 G; 387 T; 0 U; 0 Other;

Query Match 43.0%; Score 651; DB 9; Length 1804;
Best Local Similarity 71.5%; Pred. No. 3.9e-152;
Matches 871; Conservative 0; Mismatches 345; Indels 3; Gaps 1;

QY 267 GCCTTATGCCAGGGGCTGCGACCACTCTCCCGCGGCTCGGGAGGCCACCAT 326
Db 30 GCGCTACTCAACCGTCAGCTCCCGCAGTCTCGCTCGCTCGCTCGCTCGCTCGCT 89
QY 327 CGAGTCCACACACGTCGATCTCAGATGTCAGAGGACTCGTGACGCTGAACCACTACAA 386
Db 90 GGAGTCTCACACGCTCTCATCACGGGTATGCAGGACTGTGTGCACCTGAATCAGTATAC 149
QY 387 GCTGCAGAGTGAGATTGGCAAGGGTGCCTAOGTGTGTGTGAGGCTGGCTTACAACAAAG 446
Db 150 CCGTGAAGGATGAAATTGGAAGGGCTCTATGTTGTCTCAAGTTGGCTTACAATGAAA 209
QY 447 TGAAGACAGACACTATGCAATGAAAGTCTTTCCAAAAGAAAGTTACTGAAGCAGTATGG 506
Db 210 TGCAATATCTACTATGCAATGAAAGTCTGTCCAAAAGAAAGTGTATCCGGCAGCGCGG 269
QY 507 CTTTCCAGTCGCGCTCCCGGAGAGGGTCCAGGCTCCCGAGGAGGACCAAGCA 566
Db 270 CTTTCCAGTCGCGCTCCACCGGAGGACCGCGCAGCTCTCTGAGGCTGCATCCAGCC 329
QY 567 GCTGTCGCCCTGGAGCGGGTGTACAGGAGATTGCCATCTCTGAAAGAGCTGGACACGT 626
Db 330 CAGGGGCCCATTTGAGCAGGTGTACAGGAAATGCCATCTCTCAAGAGCTGGACCAACC 389
QY 627 GAATGTGTCAAACTGATCGAGGTCCTGGATGACCCAGCTGAGGACCACTCTATTGGT 686
Db 390 CAATGTGTGAAGCTGGTGGAGGTCCTGGATGACCCCAATGAGGACCATCTGTATCATGT 449
QY 687 GTTTGACCTCTGAGAAAGGGCGCGCTCATGGAAGTGCCCTGTGACAAAGCCCTTCTCGGA 746
Db 450 GTTCGAATGTTGTAACCAAGGCGCGTGTGATGGAAGTGCCCACTCAACCACTCTCTGA 509
QY 747 GGAGCAAGCTCGCTTACCTGCGGAGCTCATCTCTGGGCTCTGAGTACTTGCATGCCA 806
Db 510 AGACGAGCGCGCTTCTACTTCCAGGATCTGATCAAGGCGCATCGAGTACTTACACTACCA 569
QY 807 GAAGATCGTCCACAGGAGACATCAAGCCATCAACCTGCTCTCTGGGGATGATGGCAGCT 866
Db 570 GAAGATCATCCCGTGACATCAACCTTCAACCTCTCTGTCGAGAAAGATGGGCAAT 629
QY 867 GAAGATCGCGGACTTTGGCGCTCAGCAACCAAGTTTGGGGGAAACGACGCTGAGCTGTCCAG 926
Db 630 CAAGATCGTGACTTTGGTGTGAGCAATGAATTCAGGGCAGTGAACGCGCTCTCTCCAA 689
QY 927 CACGGCGGAACCCAGCATTCATGGCCCCCGAGGCCATTTCTGATTCGGGCCAGAGCTT 986
Db 690 CACCGTGGGACGCGCGCTTCTATGGCACCGCGCTCTCTGAGAGCCCGCAAGATCTT 749
QY 987 CAGTGGGAAGGCTTGGATGTATGGGCACCTGGGCTGACGTTGTACTGCTTGTCTATGG 1046
Db 750 CTCTGGGAAGGCTTGGATGTTTGGGCCATGGGTGTGACACTATCTACTTGTCTTTGG 809
QY 1047 GAATGCGCAATTCATGACGATTTTCACTCGGCCCTTCCAGAGGAATCAAGAAATGAGCC 1106
Db 810 CCAATGCGCAATTCATGACGAGCGGATCATGTGTTTACAGATGAAGTCAAGAGTCAAGC 869
QY 1107 CGTGGTGTCTTCAGGAGCCAGAAATCAGCGAGGAGCTCAAGGACCTGATCCTCAAGAT 1166
Db 870 CCTGGAATTTCCAGACCGCCCGACATAGCTGAGGACTTGAAGGACCTGATCAACCCGTAT 929


```

Db 1226 CTCGGGAAGGCGTTGGATGTTGGGCCATGGGTGACACTATACTGCTTGTCTTTGG 1285
QY 1047 GAAGTGCCTCATTCGACGATTTTCATCCTGGCCCTCCACAGGAAGATCAAGAATGAGCC 1106
Db 1286 CCAGTGCCTCATTCGACGAGCGGATCATGTGTTTACACAGTAAAGATCAAGAGTCAGGC 1345
QY 1107 CGTGTGTTTCTGTAGGAGCCAGAAATCAGCCAGGAGCTCAAGGACCTGATCCTGAAGAT 1166
Db 1346 COTGGAATTTCCAGACCAGCCCGACATAGCTGAGGACTTGAAGGACCTGATCACCCGTAT 1405
QY 1167 GTTAGACAAGATCCCGAGACGAGAAATGGGGTGCAGACATCAAGTTGCACCCCTTGGGT 1226
Db 1406 GCTGACACAGAAACCCCGAGTCAGGATCGTGTGCGGAATCAAGCTGCACCCCTGSGT 1465
QY 1227 GACCAAGAACGGGGAGGAGCCCTTCTTTCGAGGAGGAGCAGCTGACGCGTGTGTGAGGT 1286
Db 1466 CACGAGGCATGGGGCGGAGCCGTTGCCGTGCGAGGATGAGAACTGCACGCTGTCGAAGT 1525
QY 1287 GACAGAGGGGAGGTTAAGAACTCAGTCAGGCTCATCCACAGCTGGACACCGTGATCCT 1346
Db 1526 GACTGAAGAGGAGGTCGAGAACTCAGTCAAAACATTCACAGCTTGGCAACCGTGATCCT 1585
QY 1347 CGTGAAGTCCATGCTGAGGAAGCGTTCCCTTTGGGAACCCGTTTGAGCCCCCAGGCACGGAG 1406
Db 1586 GGTGAGACCATGATACGTAAACGCTCTTTGGGAACCCATTTCGAGGC---AGCCGGCG 1642
QY 1407 GGAAGAGCGATCATGTCTGCTCCAGGAAACCTACTGTGGAAGAAAGGTTTGGTGAAGG 1466
Db 1643 GGAGGAACCGTCACTGTACGCGCTGGAACCTTGTCTACCAAAAAACCAACGAGGGAATG 1702
QY 1467 GGGCAAGAGCCCGAGGCTC 1485
Db 1703 TGAGTCCCTGTCTGAGCTC 1721

```

Search completed: July 25, 2004, 07:15:13
 Job time : 902 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 24, 2004, 22:25:04 ; Search time 9534 Seconds
(without alignments)
6887.424 Million cell updates/sec

Title: US-10-690-617-1_COPY_173_1687

Perfect score: 1515

Sequence: 1 atggaggggggtccagctgt.....agggaagcagggctgcaccc 1515

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

1: gb.ba.*

2: gb.htg.*

3: gb.in.*

4: gb.om.*

5: gb.ov.*

6: gb.pat.*

7: gb.ph.*

8: gb.pl.*

9: gb.pr.*

10: gb.ro.*

11: gb.sts.*

12: gb.sy.*

13: gb.un.*

14: gb.vi.*

15: em.ba.*

16: em.fun.*

17: em.hum.*

18: em.in.*

19: em.mu.*

20: em.om.*

21: em.or.*

22: em.ov.*

23: em.pat.*

24: em.ph.*

25: em.pl.*

26: em.ro.*

27: em.sts.*

28: em.un.*

29: em.vi.*

30: em.htg.hum.*

31: em.htg.inv.*

32: em.htg.other.*

33: em.htg.mus.*

34: em.htg.pln.*

35: em.htg.rod.*

36: em.htg.mam.*

37: em.htg.virt.*

38: em.sy.*

39: em.htgo.hum.*

40: em.htgo.mus.*

41: em.htgo.other.*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1515	100.0	2190	6	AR221281 Sequence
2	1515	100.0	2190	6	AX406674 Sequence
3	1511.8	99.8	3529	9	AF425301 Homo sapi
4	1511.8	99.8	3536	9	AF425232 Homo sapi
5	1511.8	99.8	3575	9	BC043487 Homo sapi
6	1511.8	99.8	3583	9	AL136576 Homo sapi
7	1510.2	99.7	1937	6	AX746188 Sequence
8	1490.6	98.4	1683	6	AX455763 Sequence
9	1490.6	98.4	2711	6	AX455761 Sequence
10	1458.8	96.3	1542	6	AX166519 Sequence
11	1457.2	96.2	3501	6	AX179641 Sequence
12	1321.2	87.2	2469	9	BC031647 Homo sapi
13	1241.4	81.9	3435	10	BC017529 Mus muscu
14	1238.2	81.7	3429	10	S83194 Ca2+/calmod
15	1236.6	81.6	3411	10	AB023658 Rattus no
16	1235	81.5	3406	10	RATGCDP L42810 Rattus norv
17	1224.6	80.8	3431	10	AF117384 Mus muscu
18	685	45.2	3898	9	AK128601 Homo sapi
19	661.8	43.7	1827	9	AB081336 Homo sapi
20	661.8	43.7	2011	9	BC026060 Homo sapi
21	661.8	43.7	5324	9	AF321387 Homo sapi
22	660.2	43.6	2545	6	BD22105 Cell sign
23	659.4	43.5	1572	5	AB083001 Xenopus l
24	657	43.4	1720	9	AK024748 Homo sapi
25	651.2	43.0	4450	9	AB018330 Homo sapi
26	651	43.0	1804	9	AF101264 Homo sapi
27	651	43.0	2975	9	AF287631 Homo sapi
28	651	43.0	5611	9	AF287630 Homo sapi
29	648.2	42.8	2661	10	BC023103 Mus muscu
30	641.4	42.3	1767	6	AX056421 Sequence
31	641.4	42.3	2772	9	AF140507 Homo sapi
32	633.2	41.8	2132	10	AB018081 Rattus no
33	628.4	41.5	1866	10	AF453383 Mus muscu
34	526	34.7	1698	9	AB081337 Homo sapi
35	526	34.7	2602	9	AF321385 Homo sapi
36	526	34.7	5195	9	AF321388 Homo sapi
37	526	34.7	5238	9	AF321386 Homo sapi
38	515.6	34.0	4763	10	AK122370 Mus muscu
39	455.6	30.1	3814	9	AL834322 Homo sapi
40	376.6	24.9	1080	9	BC000318 Homo sapi
41	361	23.8	29629	6	AR221282 Sequence
42	361	23.8	29629	6	AX406676 Sequence
43	361	23.8	158414	9	AC005940 Homo sapi
44	354.4	23.4	432	6	AX775531 Sequence
45	305.6	20.2	1159	9	AF091074 Homo sapi

ALIGNMENTS

RESULT 1
AR221281 LOCUS AR221281 Sequence 1 from patent US 6426206, linear PAT 26-SEP-2002
DEFINITION AR221281 Sequence 1 from patent US 6426206, linear PAT 26-SEP-2002
ACCESSION AR221281
VERSION AR221281.1 GI:23328248
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2190)
AUTHORS Wei, M.-H., Di Francesco, V. and Beasley, E.M.
TITLE Isolated human kinase proteins, nucleic acid molecules encoding human kinase proteins, and uses thereof
JOURNAL Patent: US 6426206-A 1 30-JUL-2002;

FEATURES		Location/Qualifiers
source		1..2190
		/organism="unknown"
		/mol_type="genomic DNA"
ORIGIN		
Query Match		100.0%; Score 1515; DB 6; Length 2190;
Best Local Similarity		100.0%; Pred. No. 0;
Matches 1515; Conservative		0; Mismatches 0; Indels 0; Gaps 0;
QY	1	ATGGAGGGGGTCCAGCTGCTGCTGCCAGGATCCTCGGCAGAGCTGGTAGACGGGTG 60
Db	173	ATGGAGGGGGTCCAGCTGCTGCTGCCAGATCCTCGGCAGAGCTGGTAGACGGGTG 232
QY	61	GCAGCCATCGATGTGACTCACTTGGAGAGGCGAGATGGTGCCCGCAGAGCCCTACTAGAAAC 120
Db	233	GCAGCCATCGATGTGACTCACTTGGAGAGGCGAGATGGTGCCCGCAGAGCCCTACTAGAAAC 292
QY	121	GGTGTGGACCCCGCCACACAGGGGCCAGAGCTGCCTCTGTGATCCCTGSCAGTACTTCAAGA 180
Db	293	GGTGTGGACCCCGCCACACAGGGGCCAGAGCTGCCTCTGTGATCCCTGSCAGTACTTCAAGA 352
QY	181	CTGCTCCAGCCCGGCCCTAGCCTCTCAGCCAGGAAGCTTTCCCTACAGGAGCGGCCAGCA 240
Db	353	CTGCTCCAGCCCGGCCCTAGCCTCTCAGCCAGGAAGCTTTCCCTACAGGAGCGGCCAGCA 412
QY	241	GGAGGCTATCTGGAGGCGCGGCTGGCCCTTATGCCACGGGGCCCTGCCAGCCACATCTCC 300
Db	413	GGAGGCTATCTGGAGGCGCGGCTGGCCCTTATGCCACGGGGCCCTGCCAGCCACATCTCC 472
QY	301	CCCGGGCCTGGCGAGGCCACCATCGAGTCCACACAGCTGGCCATCTCAGATGCAGAG 360
Db	473	CCCGGGCCTGGCGAGGCCACCATCGAGTCCACACAGCTGGCCATCTCAGATGCAGAG 532
QY	361	GACTGCGTGCAGCTGAACCACTACAAGCTGCAGAGTCAGATTGGCAAGGGTGCCACGGT 420
Db	533	GACTGCGTGCAGCTGAACCACTACAAGCTGCAGAGTCAGATTGGCAAGGGTGCCACGGT 592
QY	421	GTGCTGAGGCTGGCCTACACGAAAGTGAAGACAGACACTATGCAATGAAAGTCCCTTCC 480
Db	593	GTGCTGAGGCTGGCCTACACGAAAGTGAAGACAGACACTATGCAATGAAAGTCCCTTCC 652
QY	481	AAAAAGAGTACTGAGCAGTATGGCTTTCCAGTCGCCCTCCCGCAGAGGGTCCCGAG 540
Db	653	AAAAAGAGTACTGAGCAGTATGGCTTTCCAGTCGCCCTCCCGCAGAGGGTCCCGAG 712
QY	541	GCTGCCCAGGAGGACACAGCAAGCAGCTGTGCCCTGGAGCGGGGTGTACAGAGAGATT 600
Db	713	GCTGCCCAGGAGGACACAGCAAGCAGCTGTGCCCTGGAGCGGGGTGTACAGAGAGATT 772
QY	601	GCCATCTGAAGAAGCTGACCCAGTGAATGGTGTCAAACTGATCGAGGTCTGGATGAC 660
Db	773	GCCATCTGAAGAAGCTGACCCAGTGAATGGTGTCAAACTGATCGAGGTCTGGATGAC 832
QY	661	CCAGCTGAGGACACCTCTATTGGTGTTCACCTCTGAGAAAGGGGCCCGTCAATGAA 720
Db	833	CCAGCTGAGGACACCTCTATTGGTGTTCACCTCTGAGAAAGGGGCCCGTCAATGAA 892
QY	721	GTGCCCTGTGACAAAGCCCTTCTCGGAGGAGCAAGCTCGCCCTTACCTCGCGGAGCGTCATC 780
Db	893	GTGCCCTGTGACAAAGCCCTTCTCGGAGGAGCAAGCTCGCCCTTACCTCGCGGAGCGTCATC 952
QY	781	CTGGGCCCTCAGTACTTGGCATCGCCAGAAGATCGTCCACAGGGACATCAAGCCATCCAAC 840
Db	953	CTGGGCCCTCAGTACTTGGCATCGCCAGAAGATCGTCCACAGGGACATCAAGCCATCCAAC 1012
QY	841	CTGCTCTCTGGGGATGATGGCAGTGAAGATCCCGACTTTGGCGTCAGCAACACAGTTT 900
Db	1013	CTGCTCTCTGGGGATGATGGCAGTGAAGATCCCGACTTTGGCGTCAGCAACACAGTTT 1072
QY	901	GAGGGGAACGCTCAGCTGTCCAGCACGGCGGGAAACCCAGCATTCATGGCCCCCGAG 960
Db	1073	GAGGGGAACGCTCAGCTGTCCAGCACGGCGGGAAACCCAGCATTCATGGCCCCCGAG 1132

QY	961	GCCATTTCTGATTCGCGCCAGAGCTTCAGTGGGAAGCCCTTGGATGTATGGGCCACTGGC 1020
Db	1133	GCCATTTCTGATTCGCGCCAGAGCTTCAGTGGGAAGCCCTTGGATGTATGGGCCACTGGC 1192
QY	1021	GTCAAGTTGACTGCTTTGCTATGGGAAGTGCCTATTCAGAGATTCATCGATTTTCATCTGGCC 1080
Db	1193	GTCAAGTTGACTGCTTTGCTATGGGAAGTGCCTATTCAGAGATTCATCGATTTTCATCTGGCC 1252
QY	1081	CTCCACAGGAAGATCAAGAATGAGCCCGTGGTGTTCCTCAGGAGCCAGAAATCAGCGAG 1140
Db	1253	CTCCACAGGAAGATCAAGAATGAGCCCGTGGTGTTCCTCAGGAGCCAGAAATCAGCGAG 1312
QY	1141	GAGCTCAAGGACCTGATTCCTGAAGATGTTTACAAGAATCCCGAGAGCAGAAATTTGGGGTG 1200
Db	1313	GAGCTCAAGGACCTGATTCCTGAAGATGTTTACAAGAATCCCGAGAGCAGAAATTTGGGGTG 1372
QY	1201	CCAGACATCAAGTTGACACCTTGGGTGACCAAGAACGGGAGGAGCCCTTCCTTCGGAG 1260
Db	1373	CCAGACATCAAGTTGACACCTTGGGTGACCAAGAACGGGAGGAGCCCTTCCTTCGGAG 1432
QY	1261	GAGGAGCACTGCAGCGCTGGTGGAGGTGACAGAGGGGAGGTTAAGAACTCAGTCAGGCTC 1320
Db	1433	GAGGAGCACTGCAGCGCTGGTGGAGGTGACAGAGGGGAGGTTAAGAACTCAGTCAGGCTC 1492
QY	1321	ATCCACAGCTGGACACCGGTGATCCTGGTGAAGTCCATGTCTGAGGAAGCGTTCTTTGGG 1380
Db	1493	ATCCACAGCTGGACACCGGTGATCCTGGTGAAGTCCATGTCTGAGGAAGCGTTCTTTGGG 1552
QY	1381	AACCCGTTTTCAGCCCCCAGGCAAGGAGGAGAGCGATCCATGTCTCTCCAGGAAACCTA 1440
Db	1553	AACCCGTTTTCAGCCCCCAGGCAAGGAGGAGAGCGATCCATGTCTCTCCAGGAAACCTA 1612
QY	1441	CTGCTGAAAGAGGGTTTGGTGAAGGGGCAAGAGCCAGAGCTCCCGCGCTCCAGGAA 1500
Db	1613	CTGCTGAAAGAGGGTTTGGTGAAGGGGCAAGAGCCAGAGCTCCCGCGCTCCAGGAA 1672
QY	1501	GACGAGCTGCATCC 1515
Db	1673	GACGAGCTGCATCC 1687

RESULT 2

AX406674

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

ORIGIN

AX406674

Sequence 1 from Patent WO0224920.

AX406674

AX406674.1 GI:21439624

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1

Beasley E.M., Wei, M.H., Bonazzi, V.R., Sanders, R.C. and di Francesco, V.C.

Isolated human kinase proteins, nucleic acid molecules encoding human kinase proteins, and uses thereof

Patent: WO 0224920-A 1 28-MAR-2002; PE Corporation (NY) (US)

Location/Qualifiers

1..2190

/organism="Homo sapiens"

/mol_type="unassigned DNA"

/db_xref="taxon:9606"

Query Match

Best Local Similarity

Matches 1515; Conservative

100.0%; Score 1515; DB 6; Length 2190;

100.0%; Pred. No. 0;

0; Mismatches 0; Indels 0; Gaps 0;

QY

1

ATGAGGGGGGTCCAGCTGCTGCTGCCAGGATCCTCGGCAGAGCTGGTAGACGGGTG 60

Db 173 ATGAGGGGGTCCAGCTGTCTGTGCTCCAGGATCCCTCGGCGAGAGCTGGTAGAACGGGTG 232
QY 61 GCAGCCATCGATGTGACTCACTTGGAGGAGGAGATGGTGGCCCGAGAGCTTACTAGAAAC 120
Db 233 GCAGCCATCGATGTGACTCACTTGGAGGAGGAGATGGTGGCCCGAGAGCTTACTAGAAAC 292
QY 121 GGTGTGGACCCCGACACCGGCGCAGAGCTGCTCTGTGATCCCTGGCGAGTACTTCAAGA 180
Db 293 GGTGTGGACCCCGACACCGGCGCAGAGCTGCTCTGTGATCCCTGGCGAGTACTTCAAGA 352
QY 181 CTGCTCCACAGCGCGCTAGCTCTCAGCCAGGAGAGCTTTCCTACAGAGGCGCCAGCA 240
Db 353 CTGCTCCACAGCGCGCTAGCTCTCAGCCAGGAGAGCTTTCCTACAGAGGCGCCAGCA 412
QY 241 GGAAGCTACTCTGAGGCGCAGGCTGGGCCCTTATGCCACGGGCGCTGCCAGCCACATCTCC 300
Db 413 GGAAGCTACTCTGAGGCGCAGGCTGGGCCCTTATGCCACGGGCGCTGCCAGCCACATCTCC 472
QY 301 CCCGGGCTTGGCGAGGCGCCACCATCGAGTCCACCGAGTGGCCATCTCAGATGAGAG 360
Db 473 CCCGGGCTTGGCGAGGCGCCACCATCGAGTCCACCGAGTGGCCATCTCAGATGAGAG 532
QY 361 GACTGCGTGCAGCTGAACAGTACAACTGCAAGCTGCAAGTGGCAAGGCTGCCCTACCGT 420
Db 533 GACTGCGTGCAGCTGAACAGTACAACTGCAAGTGGCAAGGCTGCCCTACCGT 592
QY 421 GTGTGAGGCTGGCTTACAAAGAAAGTGAAGACAGACACTATGCAATGAAAGTCTTTTCC 480
Db 593 GTGTGAGGCTGGCTTACAAAGAAAGTGAAGACAGACACTATGCAATGAAAGTCTTTTCC 652
QY 481 AAAAAGAGTTACTGAAGCAGTATGGCTTTCCAGTGGCCCTCCCGAGAGGTCCTCAG 540
Db 712 AAAAAGAGTTACTGAAGCAGTATGGCTTTCCAGTGGCCCTCCCGAGAGGTCCTCAG 712
QY 541 GCTGCCAGGAGGAGCAGCAGAGCTGCTGCCCTGGAGCGGCTGTACCAAGAGATT 600
Db 772 GCTGCCAGGAGGAGCAGCAGAGCTGCTGCCCTGGAGCGGCTGTACCAAGAGATT 772
QY 601 GCCATCTGGAAGAGCTGGACCAAGTGAATGTGTCAAAGTATCGAGTCTCTGGATGAC 660
Db 773 GCCATCTGGAAGAGCTGGACCAAGTGAATGTGTCAAAGTATCGAGTCTCTGGATGAC 832
QY 661 CCAGTGGAGCAACCTTATTTGTTGTTGACCTCTGAGAAAGGCGCGTCTATGAA 720
Db 833 CCAGTGGAGCAACCTTATTTGTTGTTGACCTCTGAGAAAGGCGCGTCTATGAA 892
QY 721 GTGCCCTGTGCAAGCCCTTCTCGAGGAGCAAGCTCGCCCTTACCTCGGCGAGCTCATC 780
Db 893 GTGCCCTGTGCAAGCCCTTCTCGAGGAGCAAGCTCGCCCTTACCTCGGCGAGCTCATC 952
QY 781 GTGGCCCTCGAGTACTTGCACTGCGAGAGATGCTTCCACAGGACATCAAGCCATCCAAC 840
Db 953 GTGGCCCTCGAGTACTTGCACTGCGAGAGATGCTTCCACAGGACATCAAGCCATCCAAC 1012
QY 841 CTGCTCTGGGGGATGATGGGACGTGAGATCGCGACTTTGGCGTCAGCAACAGTTT 900
Db 1013 CTGCTCTGGGGGATGATGGGACGTGAGATCGCGACTTTGGCGTCAGCAACAGTTT 1072
QY 901 GAGGGAAACGAGCTCAGCTGTCCAGCAGCGGGGAACCCAGCATTCATGGCCCCGAG 960
Db 1073 GAGGGAAACGAGCTCAGCTGTCCAGCAGCGGGGAACCCAGCATTCATGGCCCCGAG 1132
QY 961 GCCATTTCTGATTCGGGCGCAGGCTTCAAGTGGGAAGGCGCTTGGATGATGGGCCACTGGC 1020
Db 1133 GCCATTTCTGATTCGGGCGCAGGCTTCAAGTGGGAAGGCGCTTGGATGATGGGCCACTGGC 1192
QY 1021 GTCAGTTGTACTGCTTTGTCTATGGGAAGTGGCCCATTCATCGAGATTTTATCTCGCC 1080
Db 1193 GTCAGTTGTACTGCTTTGTCTATGGGAAGTGGCCCATTCATCGAGATTTTATCTCGCC 1252
QY 1081 CTCACAGGAAGATCAAGAAATGAGCCCTGTGTGTTTCTCTGAGGAGCCAGAAATCAGCGAG 1140
Db 1253 CTCACAGGAAGATCAAGAAATGAGCCCTGTGTGTTTCTCTGAGGAGCCAGAAATCAGCGAG 1312

QY 1141 GAGCTCAAGGAGCTGATCCTGAAGATGTTAGACAAGATCCCGACGAGAATTTGGGGTG 1200
Db 1313 GAGCTCAAGGAGCTGATCCTGAAGATGTTAGACAAGATCCCGACGAGAATTTGGGGTG 1372
QY 1201 CCAGACATCAAGTTGCAACCTTTGGGTGACCAAGAACGGGGAGGAGCCCTTCTCTGGAG 1260
Db 1373 CCAGACATCAAGTTGCAACCTTTGGGTGACCAAGAACGGGGAGGAGCCCTTCTCTGGAG 1432
QY 1261 GAGGAGCACTCCAGCGTGGTGAGGTGACAGAGGGGGAGGTTAAAGACTCAGTCAGGCTC 1320
Db 1433 GAGGAGCACTCCAGCGTGGTGAGGTGACAGAGGGGGAGGTTAAAGACTCAGTCAGGCTC 1492
QY 1321 ATCCCCAGCTGGACCAACCGTGTATCCTGTGAAGTCCATGCTGAGGAAGGCTTCTTTGGG 1380
Db 1493 ATCCCCAGCTGGACCAACCGTGTATCCTGTGAAGTCCATGCTGAGGAAGGCTTCTTTGGG 1552
QY 1381 AACCCGTTTGAAGGAGGTTTGGTGAAGGGGCAAGAGCCAGAGCTCCCGGCGTCCAGGA 1440
Db 1553 AACCCGTTTGAAGGAGGTTTGGTGAAGGGGCAAGAGCCAGAGCTCCCGGCGTCCAGGA 1612
QY 1441 CTGTGGAAGAGGTTTGGTGAAGGGGCAAGAGCCAGAGCTCCCGGCGTCCAGGA 1500
Db 1613 CTGTGGAAGAGGTTTGGTGAAGGGGCAAGAGCCAGAGCTCCCGGCGTCCAGGA 1672
QY 1501 GACGAGGCTGCATCC 1515
Db 1673 GACGAGGCTGCATCC 1687

RESULT 3
AF425301

LOCUS 3529 bp mRNA linear PRI 04-OCT-2002
DEFINITION Homo sapiens CAMKK alpha protein (CAMKK1) mRNA, complete cds.
ACCESSION AF425301
VERSION AF425301.1 GI:23499315

KEYWORDS
SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 3529)
AUTHORS Harvey,M., Carra,S., Tascadda,F. and Barden,N.
TITLE Characterization of human CamKK alpha gene structure
JOURNAL Unpublished

REFERENCE

2 (bases 1 to 3529)
AUTHORS Harvey,M., Carra,S., Tascadda,F. and Barden,N.
TITLE Direct Submission

JOURNAL

Submitted (27-SEP-2001) Neuroscience, CHUL, 2705 Blvd. Laurier,
Ste-Foy, Quebec G1V4G2, Canada

FEATURES

source

1..3529
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="17"
/map="17p13"
/tissue_type="brain"

gene

1..3529
/gene="CAMKK1"
95..1612
/note="protein kinase"

CDS

/codon_start=1
/product="CAMKK alpha protein"
/protein_id="AA37387.1"
/db_xref="GI:23499316"

/translation="MEGPAVCCQDPRAELVERVAAIDVTHLEADGGPEPTRNGVDP
PPRAASVLPGSTRLPPARPSLSARKLSIQRPAAGVYLAQAGPVATGPASHISPR
AWRPPTTSHHVAISDAEDCVQNOYKLOSEIGKYGTVRLAYNESEDHRYAMKVL
KKLLKQYGFRRPPRPGSAQGPAPKQLPLERVYQETAILKLDHNVKLVLEVL
DDPAEDNLXILVFDLLRQPVMEVPCDKPFSEBQARLYLDVILGLEVYHCQKIVHRDI
KPSNLLLDGDDGHVKIADFGVSNQFEQNDLAQLSSSTAGTAPFAPMAPEAIDSQGSFSGKAL

DVWATGVTLYCFVYKCPFFIDDFILALHXRKNEPVPVFPPEPISBELKDLILKMDIK
NPTRIGVDPDKLHPWTKNEPPLSEEBHCSVVEVEEVRNVRLLIPSWTIVILV
KSMRLKRSFQNPFPQARRERESMSAPGNLLVKEFGEGGKSPELPGVQDEAAS"

ORIGIN

QY	1	ATGAGGGGGGTCCAGCTGCTGCTGCGCAGGATCCTCGGGCAGAGCTGGTAGAACGGGTG	60
Db	95	ATGAGGGGGGTCCAGCTGCTGCTGCGCAGGATCCTCGGGCAGAGCTGGTAGAACGGGTG	154
QY	61	GCAGCCATCATGATGACTCACTTGGAGGAGCAGATGGTGGCCAGAGCCCTACTAGAAAC	120
Db	155	GCAGCCATCATGATGACTCACTTGGAGGAGCAGATGGTGGCCAGAGCCCTACTAGAAAC	214
QY	121	GGTGTGACCCCCACACAGGGGCCAGAGCTGCTGCTGATCCCTGGCAGTACTTCAAGA	180
Db	215	GGTGTGACCCCCACACAGGGGCCAGAGCTGCTGCTGATCCCTGGCAGTACTTCAAGA	274
QY	181	CTGCTCCAGCCCGGCTAGCTCTCAGCCAGGAGCTTTCCCTACAGGAGCGGCCAGCA	240
Db	275	CTGCTCCAGCCCGGCTAGCTCTCAGCCAGGAGCTTTCCCTACAGGAGCGGCCAGCA	334
QY	241	GGAGCTATCTGGAGGCGCAGGCTGGGCTTATGCCAGGGGCTGCCAGCACATCTCC	300
Db	335	GGAGCTATCTGGAGGCGCAGGCTGGGCTTATGCCAGGGGCTGCCAGCACATCTCC	394
QY	301	CCCCGGGCTGGGGAGGCCACCATGAGTCCACCAAGCTGGCCATCTCAGATGCAGAG	360
Db	395	CCCCGGGCTGGGGAGGCCACCATGAGTCCACCAAGCTGGCCATCTCAGATGCAGAG	454
QY	361	GACTGCGTGCAGCTGAACAGCTACAAGCTGAGAGTGGCAAGGGTGCCTACGGT	420
Db	455	GACTGCGTGCAGCTGAACAGCTACAAGCTGAGAGTGGCAAGGGTGCCTACGGT	514
QY	421	GTGTGAGGCTGGGCTACAAAGTGAAGACAGACATATGCAATGAAGTCTCTTCC	480
Db	515	GTGTGAGGCTGGGCTACAAAGTGAAGACAGACATATGCAATGAAGTCTCTTCC	574
QY	481	AAAGAGAGTACTCAAGCAGTATGGCTTCCAGTGGCTCCCGGAGAGGTCCTCAG	540
Db	575	AAAGAGAGTACTCAAGCAGTATGGCTTCCAGTGGCTCCCGGAGAGGTCCTCAG	634
QY	541	GCTGCCAGGAGGACCAAGCAGCTGCTGCCCTGGAGCGGGTGTACAGGAGATT	600
Db	635	GCTGCCAGGAGGACCAAGCAGCTGCTGCCCTGGAGCGGGTGTACAGGAGATT	694
QY	601	GCCATCTGAAGAGCTGGACCAAGTGAATGTGTCAAACTGATCGAGTCTCGATGAC	660
Db	695	GCCATCTGAAGAGCTGGACCAAGTGAATGTGTCAAACTGATCGAGTCTCGATGAC	754
QY	661	CCAGCTGAGGACCACTCTATTGGTGTTCACCTCTGAGAAAGGGCCCGTCAATGAA	720
Db	755	CCAGCTGAGGACCACTCTATTGGTGTTCACCTCTGAGAAAGGGCCCGTCAATGAA	814
QY	721	GTGCCCTGTGACAAAGCCCTCTCGAGGAGCAAGCTCGCCCTTACCTCGCGGAGCTCATC	780
Db	815	GTGCCCTGTGACAAAGCCCTCTCGAGGAGCAAGCTCGCCCTTACCTCGCGGAGCTCATC	874
QY	781	CTGGGCTCGAGTACTTGCATGTCAGAGATGCTCCAGGAGCATCAAGCCATCCAC	840
Db	875	CTGGGCTCGAGTACTTGCATGTCAGAGATGCTCCAGGAGCATCAAGCCATCCAC	934
QY	841	CTGCTCTGGGGGATGATGGCAGCTGAAGATCCCGAGCTTTGGCGTCAGCAACAGATT	900
Db	935	CTGCTCTGGGGGATGATGGCAGCTGAAGATCCCGAGCTTTGGCGTCAGCAACAGATT	994
QY	901	GAGGGGAACAGCTGATGCTGTCAGACGCGCGGGAACCCAGCATTCATGGCCCCCGAG	960
Db	995	GAGGGGAACAGCTGATGCTGTCAGACGCGCGGGAACCCAGCATTCATGGCCCCCGAG	1054

QY	961	GCCATTTCTGATTCGGCCAGAGCTTCAGTGGGAGGSCCTTGGATGTATGGGCCACTGGC	1020
Db	1055	GCCATTTCTGATTCGGCCAGAGCTTCAGTGGGAGGSCCTTGGATGTATGGGCCACTGGC	1114
QY	1021	GTACAGTTGTACTGCTTTGTCTATGGGAGTGCCTTATTCATGACGATTTTCATCTGGCC	1080
Db	1115	GTACAGTTGTACTGCTTTGTCTATGGGAGTGCCTTATTCATGACGATTTTCATCTGGCC	1174
QY	1081	CTCCACAGGAAGATCAAGAAATGAGCCCGTGTGTCTTCTGAGGAGCCAGAAATCAGCGAG	1140
Db	1175	CTCCACAGGAAGATCAAGAAATGAGCCCGTGTGTCTTCTGAGGAGCCAGAAATCAGCGAG	1234
QY	1141	GAGCTCAAGAACCTGATCTTGAAGATGTTAGACAAGATCCCGAGAGAGAAATGGGGTG	1200
Db	1235	GAGCTCAAGAACCTGATCTTGAAGATGTTAGACAAGATCCCGAGAGAGAAATGGGGTG	1294
QY	1201	CCAGACATCAAGTTGACACCTTGGGTGACCAAGAACGGGAGAGGCCCTTCCCTTCGGAG	1260
Db	1295	CCAGACATCAAGTTGACACCTTGGGTGACCAAGAACGGGAGAGGCCCTTCCCTTCGGAG	1354
QY	1261	GAGGAGCACTGCAGCGTGTGGAGGTGACAGAGGGGAGGTTAAGAACTCAGTCAGGCTC	1320
Db	1355	GAGGAGCACTGCAGCGTGTGGAGGTGACAGAGGGGAGGTTAAGAACTCAGTCAGGCTC	1414
QY	1321	ATCCCGAGCTGGACCAAGCTGATCTGCTGAAGTCCATGCTGAGGAGAGGTCCTTTGGG	1380
Db	1415	ATCCCGAGCTGGACCAAGCTGATCTGCTGAAGTCCATGCTGAGGAGAGGTCCTTTGGG	1474
QY	1381	AACCCGTTTGAAGCCAGGACCGAGGAGGAGGATCCATGCTGCTCCAGGAAACCTA	1440
Db	1475	AACCCGTTTGAAGCCAGGACCGAGGAGGAGGATCCATGCTGCTCCAGGAAACCTA	1534
QY	1441	CTGTGAAAGAGGTTTGTGAAGGGGCAAGAGCCAGAGCTCCCGGGGTCAGGAA	1500
Db	1535	CTGTGAAAGAGGTTTGTGAAGGGGCAAGAGCCAGAGCTCCCGGGGTCAGGAA	1594
QY	1501	GACGAGGCTGCATCC	1515
Db	1595	GACGAGGCTGCATCC	1609

RESULT 4

LOCUS	AF425232	Homo sapiens CaMKK alpha protein mRNA, complete cds.
DEFINITION	AF425232	3536 bp mRNA linear PRI 04-OCT-2002
ACCESSION	AF425232	
VERSION	AF425232.1	GI:23499313
KEYWORDS		Homo sapiens (human)
SOURCE		Homo sapiens
ORGANISM		Homo sapiens
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS		Tascadda, F., Carra, S., Harvey, M. and Barden, N.
TITLE		Characterization of human CaMKK alpha gene structure
JOURNAL		Unpublished
REFERENCE		2 (bases 1 to 3536)
AUTHORS		Tascadda, F., Carra, S., Harvey, M. and Barden, N.
TITLE		Direct Submission
JOURNAL		Submitted (27-SEP-2001) Pharmacology, University of Modena and Reggio Emilia, Campi 183, Modena, MO 41100, Italy
FEATURES		Location/Qualifiers
source		1..3536
		/organism="Homo sapiens"
		/mol_type="mRNA"
		/db_xref="taxon:9606"
		/chromosome="17"
		/map="17p13"
		/tissue_type="brain"
		102..1619
		/note="protein kinase kinase; CaMKK 1"
		/codon_start=1
		/product="CaMKK alpha protein"
CDS		

/protein id="AAN37386.1"
/db_xref="GI:2349314"
/translation="MEGGPACVCCOPRAELVERVAADIVTHLEADGGPEPTRNGVDP
PPRARAASVPGSTSELLPAPSLARKLSLOERPAQSVLEAQAQPYATGPASHLSPR
AWERTTISHHVAISDAEDCVOLNQYKQSEIGKAGYGVVRLAYNESDRHYAMKVL
KKLJKQGFPRPFRSQAAGGPAQKLLPLERYOEIALKLLKHVNKKLLEVL
DDPNDNLYLFDLLKRGVMEVPCDPFSEEQARLYLFDVLGLYEYFCQKIVHRDI
KPSNLLGDGDKVADFGVSNQFNGDAQLSAGTAPFAPEAISDSGGFSKGL
DYATGVTLYCFVYKCPFDIDFIALHRKIKNEPVVRPEPISSEELKDLILKMLDK
NPETRIGVEDIKLHPWTKNGEPELPSSEEHCSVVEVEEVEKNSVRLIPSWTIVILV
KSMLEKRSFGNPPFQARREERSMSPGNLLVKEGFGGKSPFLPGVQEDBAAS"

Query Match 99.8%; Score 1511.8; DB 9; Length 3536;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1513; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

ORIGIN

1 ATGGAGGGGGTCCAGCTGTCTGTCGCCAGGATCCTCGGGCAGAGCTGTAGAACGGGTG 60
102 ATGGAGGGGGTCCAGCTGTCTGTCGCCAGGATCCTCGGGCAGAGCTGTAGAACGGGTG 161
61 GCAGGCATCGATGTCACTCACTTGGAGAGGCAGATGTTGGCCAGAGCCTACTAGAAC 120
162 GCAGGCATCGATGTCACTCACTTGGAGAGGCAGATGTTGGCCAGAGCCTACTAGAAC 221
121 GGTGTGGACCCCCACCCAGCGGCCAGAGCTGCTCTGTGATCCCTGGCAGTACTTCAAGA 180
222 GGTGTGGACCCCCACCCAGCGGCCAGAGCTGCTCTGTGATCCCTGGCAGTACTTCAAGA 281
181 CTGCTCCAGCCCGGCTAGCTCTCAGCCAGGAGCTTCCCTACAGAGCGGCCAGCA 240
282 CTGCTCCAGCCCGGCTAGCTCTCAGCCAGGAGCTTCCCTACAGAGCGGCCAGCA 341
241 GGAAGTATCTGAGCGCGAGCTGGGCTTATGCGAGGGGCTTCCAGCCACATCTCC 300
342 GGAAGTATCTGAGCGCGAGCTGGGCTTATGCGAGGGGCTTCCAGCCACATCTCC 401
301 CCCCGGGCTGCGGAGGCCACCATCGAGTCCCAACCGTGGCCATCTCAGATCAGAG 360
402 CCCCGGGCTGCGGAGGCCACCATCGAGTCCCAACCGTGGCCATCTCAGATCAGAG 461
361 GACTCGGTGACGTGAACAGTGAACAGTGAACAGTGAACAGTGAACAGTGAACAGTGA 420
462 GACTCGGTGACGTGAACAGTGAACAGTGAACAGTGAACAGTGAACAGTGAACAGTGA 521
421 GTGTGAGGCTGGCTTACAAAGAGTGAAGAGTGAAGAGTGAAGAGTGAAGAGTGAAG 480
522 GTGTGAGGCTGGCTTACAAAGAGTGAAGAGTGAAGAGTGAAGAGTGAAGAGTGAAG 581
481 AAAAAGAGTGAAGAGTGAAGAGTGAAGAGTGAAGAGTGAAGAGTGAAGAGTGAAGAG 540
582 AAAAAGAGTGAAGAGTGAAGAGTGAAGAGTGAAGAGTGAAGAGTGAAGAGTGAAGAG 641
541 GCTGCCAGGAGGACCGCAGAGCTGCTGCCCTGGAGCGGCTGTACAGAGAGATT 600
642 GCTGCCAGGAGGACCGCAGAGCTGCTGCCCTGGAGCGGCTGTACAGAGAGATT 701
601 GCATCTCTGAAGAGTGAAGAGTGAAGAGTGAAGAGTGAAGAGTGAAGAGTGAAGAGTGA 660
702 GCATCTCTGAAGAGTGAAGAGTGAAGAGTGAAGAGTGAAGAGTGAAGAGTGAAGAGTGA 761
661 CCAGCTGAGGACAACTCTATTGTTGTTGTTGACCTCTGAGAAAGGGCGCTCATGGAA 720
762 CCAGCTGAGGACAACTCTATTGTTGTTGTTGACCTCTGAGAAAGGGCGCTCATGGAA 821
721 GTGCCCTGTGACAAAGCCCTTCTCGAGAGGACAGTCTGCCCTACTCTGGGAGACGTATC 780
822 GTGCCCTGTGACAAAGCCCTTCTCGAGAGGACAGTCTGCCCTACTCTGGGAGACGTATC 881
781 GTGGGCCCTCGAGTACTTGCCTGCGCAGAGATGCTGCCAGGACATCAAGCCATCCCAAC 840
882 GTGGGCCCTCGAGTACTTGCCTGCGCAGAGATGCTGCCAGGACATCAAGCCATCCCAAC 941

QY 841 CTGCTCCTGGGGATGATGGGCACGTGAAGATCGCCGACTTTGGGCTCAGCAACCAAGTTT 900
DB 942 CTGCTCCTGGGGATGATGGGCACGTGAAGATCGCCGACTTTGGGCTCAGCAACCAAGTTT 1001
QY 901 GAGGGGAACGACGCTCAGCTGATCCAGCAGCGGGGAAACCCAGCATTCATATGCCCCCGAG 960
DB 1002 GAGGGGAACGACGCTCAGCTGATCCAGCAGCGGGGAAACCCAGCATTCATATGCCCCCGAG 1061
QY 961 GCATTTCTGATTCGCGGCAGAGCTTCACTGGAGAGGCTTGGATGATGGGCCACTGGC 1020
DB 1062 GCATTTCTGATTCGCGGCAGAGCTTCACTGGAGAGGCTTGGATGATGGGCCACTGGC 1121
QY 1021 GTCACTTGTACTGCTTTGTTCTATGGGAAGTCCCAATTCACTGACGATTCATCTCTGGCC 1080
DB 1122 GTCACTTGTACTGCTTTGTTCTATGGGAAGTCCCAATTCACTGACGATTCATCTCTGGCC 1181
QY 1081 CTCACAGAGAGATCAAGATGAGCCCGTGTGTTCCTGAGAGCCAGAAATCAGCGAG 1140
DB 1182 CTCACAGAGAGATCAAGATGAGCCCGTGTGTTCCTGAGAGCCAGAAATCAGCGAG 1241
QY 1141 GAGCTCAAGGACCTGATCTGTAAGATGTTAGACAAGATCCCGAGACGAGATTCGGGTG 1200
DB 1242 GAGCTCAAGGACCTGATCTGTAAGATGTTAGACAAGATCCCGAGACGAGATTCGGGTG 1301
QY 1201 CCAGACATCAAGTTGACCCCTTGGGTGACCAAGAACGGGGAGAGCCCTTCTCTCGGAG 1260
DB 1302 CCAGACATCAAGTTGACCCCTTGGGTGACCAAGAACGGGGAGAGCCCTTCTCTCGGAG 1361
QY 1261 GAGGAGCACTGAGCGTGTGTGAGTGACAGAGGGGAGGTTAAGAACTCAGTCAGGCTC 1320
DB 1362 GAGGAGCACTGAGCGTGTGTGAGTGACAGAGGGGAGGTTAAGAACTCAGTCAGGCTC 1421
QY 1321 ATCCCGACGTGACCAAGTGTATCTGTAAGTCCATGCTGAGGAAGCTTTCCTTGGG 1380
DB 1422 ATCCCGACGTGACCAAGTGTATCTGTAAGTCCATGCTGAGGAAGCTTTCCTTGGG 1481
QY 1381 AACCCGTTGAGCCCCAGGCGACGAGGGAAGAGCGATCCATGTCTCTCCAGGAACCTA 1440
DB 1482 AACCCGTTGAGCCCCAGGCGACGAGGGAAGAGCGATCCATGTCTCTCCAGGAACCTA 1541
QY 1441 CTGTGGAAGAGGGTTTGGTGAAGGGGGCAAGAGCCAGAGCTCCCGGCGTCCAGAA 1500
DB 1542 CTGTGGAAGAGGGTTTGGTGAAGGGGGCAAGAGCCAGAGCTCCCGGCGTCCAGAA 1601
QY 1501 GACGAGGCTGCATCC 1515
DB 1602 GACGAGGCTGCATCC 1616

RESULT 5
LOCUS BC043487
DEFINITION Homo sapiens calcium/calmodulin-dependent protein kinase kinase 1,
alpha, transcript variant 1, mRNA (cDNA clone MGC:49844
IMAGE:5751570), complete cds.
ACCESSION BC043487
VERSION BC043487.1 GI:27694083
KEYWORDS MGC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 3575)
AUTHORS Klausner,R.D., Collins,F.S., Wagner,L., Shennan,C.M., Schuler,G.D.,
Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
Abramson,R.D., Mullaly,S.J., Bosak,S.A., McEwan,P.J.,
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,

Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
Valladao, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahey, J., Helton, E., Kettelman, M., Maman, A., Rodriguez, S.,
Sanchez, A., Whiting, M., Maman, A., Young, A.C., Shevchenko, Y.,
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalhus, D.E.,
Schnerch, A., Schein, J.E., Jones, S.J., and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
22388257
12477932
2 (bases 1 to 3575)
Strausberg, R.
Direct Submission
Submitted (09-JAN-2003) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgapsb@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILLNL)
DNA Sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
info@bcsc.bc.ca
Steven Jones, Jennifer Aeano, Ian Bosdet, Yaron Butterfield,
Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
Leticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo
Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven
Ness, Pawan Pandoh, Anna-Lilisa Prabhu, Parvaneh Saeedi, Jacqueline
Schein, Duane Smalhus, Michael Smith, Lorraine Spence, Jeff Stott,
Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,
George Yang, Scott Zuyderduyn, Marco Marra.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/ILLNL at: <http://image.llnl.gov>
Series: IRAK Plate: 89 Row: h Column: 18
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 27437009.
FEATURES
source
Location/Qualifiers
1..3575
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:49844 IMAGE:5751570"
/tissue_type="Brain, Lung, Testis, adult, pooled whole"
/clone_lib="NIH_MGC_115"
/lab_host="DH10B"
/note="vector: pCMV-SPORT6"
1..3575
/gene="CAMK1A"
/note="synonyms: DKFZp761M0423, CAMK1A, MGC34095"
/db_xref="LocusID:84254"
125..1642
/codon_start=1
/product="calcium/calmodulin-dependent protein kinase 1
alpha, isoform a"
/protein_id="AAH43487.1"
/db_xref="GI:27694084"
/db_xref="LocusID:84254"
/translation="MEGFAVCCQDPRAELVERVAALDVTHLEADGGPPTNRGVDP
PPRAASVLPSTSRLLPAPRSLASARKLSQERPAQSLAEQAGPVATPASHISPR
AWRRPTTESHVAISDAEDCVLNQYKQSEIKGAYGVVRLAYNESEDHRYAMKVL
KKLLQYGFPPRPPRGSQAQGGPAKQLPLERVQETAILKLDHNVVVKLIEVL
DDPAEDNLVLDLRKPVNVEVPDKPFEEOARLYRDVILGLEVYHCQKI VHRDI
KPSNLILGDGHVADIGVSNQFEGNDAQLSSAGTFAEMAPEAISDSGSPGKAL
DVWATGVLVYCVYKCPFDIDFIALHRLKNEPVVFPPEPEISELKDILKMLDK
NPETRIQVDPDILKHPWTKNGEEBPLPSEEEHCVSEVTEBEVKNRSLIIPSWTIVILV
KSLMRKRSFGNPFQARRERSMSAPGNLLVKEGFGEGGKSPGLPGVQDEAAS"

misc_feature 506..1351
/note="S-TKc; Region: Serine/Threonine protein kinases,
catalytic domain"
/db_xref="CDD:smart00220"
ORIGIN
Query Match 99.8%; Score 1511.8; DB 9; Length 3575;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1513; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 ATGAGAGGGGGTCCAGCTGTCTGCTCCAGGATCTCTCGGCAGAGCTGTGTAGAACGGGTG 60
DB 125 ATGAGAGGGGGTCCAGCTGTCTGCTCCAGGATCTCTCGGCAGAGCTGTGTAGAACGGGTG 184
QY 61 GAGGCCATCATGTGACTCTTGGAGAGGAGAGAGTGTGGCCAGAGCCCTACTAGAAC 120
DB 185 GCAGCCATCATGTGACTCTTGGAGAGGAGAGTGTGGCCAGAGCCCTACTAGAAC 244
QY 121 GGTGTGACCCCCACCACCGGGCCAGAGCTCTGTGATCCCTGGCAGTACTTCAAGA 180
DB 245 GGTGTGACCCCCACCACCGGGCCAGAGCTCTGTGATCCCTGGCAGTACTTCAAGA 304
QY 181 CTGCTCCAGCCCGGCTAGCTCTCTACGCCAGGAAGCTTTCCCTACAGGAGCGCCAGCA 240
DB 305 CTGCTCCAGCCCGGCTAGCTCTCTACGCCAGGAAGCTTTCCCTACAGGAGCGCCAGCA 364
QY 241 GGAAGCTATCTGAGGCGGAGCTGGGCTTATGCCAGGGGCTGCGCAGCCATCTCC 300
DB 365 GGAAGCTATCTGAGGCGGAGCTGGGCTTATGCCAGGGGCTGCGCAGCCATCTCC 424
QY 301 CCCGGGCTTGGGAGGCGCCACCATCGAGTCCACCACCGTGGCCATCTCAGATGCGAG 360
DB 425 CCCGGGCTTGGGAGGCGCCACCATCGAGTCCACCACCGTGGCCATCTCAGATGCGAG 484
QY 361 GACTCGCTGAGCTGAACACAGTACAGCTGACAGTGTGCAAGGGTGTCTACGGT 420
DB 485 GACTCGCTGAGCTGAACACAGTACAGCTGACAGTGTGCAAGGGTGTCTACGGT 544
QY 421 GTGGTGGGCTGGCTTACAGGAAGTGAAGACAGACACTATGCAATGAAGTCTTTCC 480
DB 545 GTGGTGGGCTGGCTTACAGGAAGTGAAGACAGACACTATGCAATGAAGTCTTTCC 604
QY 481 AAAAAAGTGTACTGAAGCAGTATGGCTTTCCAGCTCGCCCTCCCGCAGAGGGTCCAG 540
DB 605 AAAAAAGTGTACTGAAGCAGTATGGCTTTCCAGCTCGCCCTCCCGCAGAGGGTCCAG 664
QY 541 GTGCCCCAGGAGGACAGCCAAAGCAGCTGCTGCCCTGGAGCGGGTGTACAGAGATT 600
DB 665 GTGCCCCAGGAGGACAGCCAAAGCAGCTGCTGCCCTGGAGCGGGTGTACAGAGATT 724
QY 601 GCCATCTGAAGAGCTGGACCGACCTGCAATGTGTCAAACTGATCGAGGTCCTGGATGAC 660
DB 725 GCCATCTGAAGAGCTGGACCGACCTGCAATGTGTCAAACTGATCGAGGTCCTGGATGAC 784
QY 661 CCAGCTGAGGACCACTCTATTGGTGTGTTGACCTCTGAGAAAGGGGCCCGTCATGAA 720
DB 785 CCAGCTGAGGACCACTCTATTGGTGTGTTGACCTCTGAGAAAGGGGCCCGTCATGAA 844
QY 721 GTGCCCTGTGAAGCCCTTCTCGGAGGAGCAAGCTCGCTTACCTCGGAGGAGTCAATC 780
DB 845 GTGCCCTGTGAAGCCCTTCTCGGAGGAGCAAGCTCGCTTACCTCGGAGGAGTCAATC 904
QY 781 CTGGGCTCGAGTACTTGGCACTGCGCAGAGATCGTCCACAGGAGCATCAAGCCATCAAC 840
DB 905 CTGGGCTCGAGTACTTGGCACTGCGCAGAGATCGTCCACAGGAGCATCAAGCCATCAAC 964
QY 841 CTGCTCTCGGGGATGATGGCAGCTGAAGATCGCGACTTTGGCGTGCAGCAACAGTTT 900
DB 965 CTGCTCTCGGGGATGATGGCAGCTGAAGATCGCGACTTTGGCGTGCAGCAACAGTTT 1024
QY 901 GAGGGGAACCAAGCTGAGTGTCCAGCAGCGGGGAACCCAGCATTCATGGCCCCCGAG 960
DB 1025 GAGGGGAACCAAGCTGAGTGTCCAGCAGCGGGGAACCCAGCATTCATGGCCCCCGAG 1084

```
QY 961 GCATTCTGATTCGGCCAGAGCTTCAGTGGGAAGCCCTTGGATGATATGGGCCACTGGC 1020
|||
Db 1085 GCCATTCTGATTCGGCCAGAGCTTCAGTGGGAAGCCCTTGGATGATATGGGCCACTGGC 1144
QY 1021 GTCACGTTGACGCTTGTCTATGGAAGTGGCCATTCATCAAGATTTCATCCTGGCC 1080
|||
Db 1145 GTCACGTTGACGCTTGTCTATGGAAGTGGCCATTCATCAAGATTTCATCCTGGCC 1204
QY 1081 CTCACAGGAAGATCAAGATGATGCGCGTGTCTTCTGAGGAGCAGAAATCAGCGAG 1140
|||
Db 1205 CTCACAGGAAGATCAAGATGATGCGCGTGTCTTCTGAGGAGCAGAAATCAGCGAG 1264
QY 1141 GAGCTCAAGACCTGATCCTGAAGATGTTAGACAAGAAATCCCGAGACGAAATGGGGTG 1200
|||
Db 1265 GAGCTCAAGACCTGATCCTGAAGATGTTAGACAAGAAATCCCGAGACGAAATGGGGTG 1324
QY 1201 CCAGACATCAAGTTGACCCCTTGGGTGACCAAGAACGGGGAGAGCCCTTCTTCGGAG 1260
|||
Db 1325 CCAGACATCAAGTTGACCCCTTGGGTGACCAAGAACGGGGAGAGCCCTTCTTCGGAG 1384
QY 1261 GAGGACATCGACGCTGTGGAGGTGACAGAGGGGAGGTTAAGAACTCAGTCAGGCTC 1320
|||
Db 1385 GAGGACATCGACGCTGTGGAGGTGACAGAGGAGGTTAAGAACTCAGTCAGGCTC 1444
QY 1321 ATCCCGAGCTGACACCGTGTCTTGGTGAAGTCCATGCTGAGGAAGCTTCTTTGGG 1380
|||
Db 1445 ATCCCGAGCTGACACCGTGTCTTGGTGAAGTCCATGCTGAGGAAGCTTCTTTGGG 1504
QY 1381 AACCGTTTGGCCCGACGACGAGGAGGAGCGATCCATGCTGCTCAGGAACCTTA 1440
|||
Db 1505 AACCGTTTGGCCCGACGACGAGGAGGAGCGATCCATGCTGCTCAGGAACCTTA 1564
QY 1441 CTGGTGAAGAGAGGTTTGTGAAGGGGCAAGAGCCAGAGCTCCCGCGTCCAGGAA 1500
|||
Db 1565 CTGGTGAAGAGAGGTTTGTGAAGGGGCAAGAGCCAGAGCTCCCGCGTCCAGGAA 1624
QY 1501 GACGAGGCTGCATCC 1515
|||
Db 1625 GACGAGGCTGCATCC 1639

RESULT 6
HSM801550 3583 bp mRNA linear PRI 10-MAR-2001
LOCUS cDNA DKFp761M0423 (from clone DKFp761M0423);
DEFINITION complete cds.
ACCESSION AL136576
VERSION AL136576.1 GI:13276654
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1
AUTHORS Wiemann,S., Weil,B., Weilenreuther,R., Gassenhuber,J., Glassl,S.,
Ansorge,W., Boecker,H., Bloecker,H., Bauersachs,S., Blum,H.,
Lauber,J., Duesterhoeft,A., Beyer,A., Koehrer,K., Strack,N.,
Mewes,H.W., Oostenwajder,B., Obermaier,B., Tampe,J., Heubner,D.,
Wambutt,R., Korn,B., Klein,M. and Poustka,A.
Toward a catalog of human genes and proteins: sequencing and
analysis of 500 novel complete protein coding human cDNAs
Genome Res. 11 (3), 422-435 (2001)
21154917
PUBMED 11230166
REFERENCE 2 (bases 1 to 3583)
AUTHORS Bloecker,H., Boecker,M., Brandt,P., Mewes,H.W., Weil,B. and
Wiemann,S.
Direct Submission
Submitted (18-JAN-2000) MIPS, Am Klopferspitz 18a, D-82152
Martinsried, GERMANY
JOURNAL Clome from S. Wiemann, Molecular Genome Analysis, German Cancer
COMMENT Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
```

sequenced by GBF (National Research Centre for Biotechnology Ltd., Braunschweig/Germany) within the cDNA sequencing consortium of the German Genome Project.
This clone is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further information about the clone and the sequencing project is available at <http://www.mips.biochem.mpg.de/proj/cDNA/>.

FEATURES

source
1..3583
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="RZPD:DKFp761M0423"
/db_xref="taxon:9606"
/map="26.0 cR from top of Chr17 linkage group"
/clone="DKFp761M0423"
/tissue_type="amygdala"
/clone_lib="761 (synonym: hamy2). Vector pSport1; host DH10B; sites NotI + SalI"
/dev_stage="adult"
1..3583
/gene="DKFp761M0423"
149..1866
/gene="DKFp761M0423"
/note="similarity to Cam-kinase alpha (Rattus norvegicus)"
/codon_start=1
/product="hypothetical protein"
/protein_id="CAB66511.1"
/db_xref="GI:13276655"
/db_xref="GOA:Q9BQH3"
/db_xref="SPTREMBL:Q9BQH3"
/translation="MEGPAVCCQSPRAELVERVAALDVTHLEADGGPEPTRNGVDP
PPRAAASVTPGSTSRLLPARELSQERFAGSYLEAQAGPYATGPASHISPR
AKWRPTIESHHVAISDAEDCVQLNQKQSEIGKAGYGVRLAYNESEDRHYAMKVL
KXLLKQYQFRRPPPPGSAQAGGPAKOLLPLRYQVITAILKLDHVNVLKLEVL
DDPAEDNLYLVFDLLRKGPVMEVPCDPFSEEQARLYRDVLIGLEYLHCOKIVHRDI
KPSNLLIGDDGHVKIADFGVSNQFESNDQAQSLSTACTPAFMAPEAISDSQSPSGKAL
DYMATGVTLVCFYGVKGPIDDFILALHRKIKNEPVVFEPEPEISEELKDLILKMLDK
NFETRIGVDPDIKLHFWVTKNBEELPFSBEHECHSVVEETEEVKNSVRLILPSPWTVILV
KSMRLRSFGNDFEFPQARREERSMSPGNLLVKEGFGGKSPFLPGVQEDFAAS"
3525..3530
/gene="DKFp761M0423"
3546
/gene="DKFp761M0423"
polya_signal
polya_site
ORIGIN

Query Match 99.8%; Score 1511.8; DB 9; Length 3583;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1513; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 ATGGAGGGGGGTCCAGCTGTCTGCTGCCAGGATCTCTCGGCGAGAGCTGTGTAGAACGGGTG 60
|||
Db 149 ATGGAGGGGGGTCCAGCTGTCTGCTGCCAGGATCTCTCGGCGAGAGCTGTGTAGAACGGGTG 208
|||
QY 61 GCAGCCATCGATGTCACTCACTTGAGGAGGAGGAGAGCTGTGGCCAGAGCTACTAGAAC 120
|||
Db 209 GCAGCCATCGATGTCACTCACTTGAGGAGGAGGAGAGCTGTGGCCAGAGCTACTAGAAC 268
|||
QY 121 GGTGTGGACCCCCACCCAGCGGCCAGAGCTGCCTCTGTGATCCCTGGCAGTACTTCAAGA 180
|||
Db 269 GGTGTGGACCCCCACCCAGCGGCCAGAGCTGCCTCTGTGATCCCTGGCAGTACTTCAAGA 328
|||
QY 181 CTGCTCCAGCCCGGCTTACGCTCTCAGCCAGGAAGCTTTCCTACAGAGGCGGCAGGA 240
|||
Db 329 CTGCTCCAGCCCGGCTTACGCTCTCAGCCAGGAAGCTTTCCTACAGAGGCGGCAGGA 388
|||
QY 241 GGAAGCTATCTGGAGGCGCAGCTGGGCTTATGCCAGGGGCGCTGCCAGCCACATCTCC 300
|||
Db 389 GGAAGCTATCTGGAGGCGCAGCTGGGCTTATGCCAGGGGCGCTGCCAGCCACATCTCC 448
|||
QY 301 CCCCGGCGCTGGCGGAGGCCACCATCGAGTCCCAACACCGTGGCCATCTCAGATCAGAG 360
|||

Db 449 CCCCCGGCTGGCGAGGCCCCACCATCGAGTCCACACAGCTGGCCATCTCAGATCGAGAG 508
Qy 361 GACTCGTGCAGCTGAACACGATACAGCTGCAGAGTGAGATTGGCAAGGGTGCCTACGGT 420
Db 509 GACTCGTGCAGCTGAACACGATACAGCTGCAGAGTGAGATTGGCAAGGGTGCCTACGGT 568
Qy 421 GTGTGAGGCTGGCTTACAAGCAAGTGAAGACAGACACTATGCAATGAAAGTCTTTTCC 480
Db 569 GTGTGAGGCTGGCTTACAAGCAAGTGAAGACAGACACTATGCAATGAAAGTCTTTTCC 628
Qy 481 AAAAAGAGTTACTGAAGCAGTATGGCTTTCCACGTGCGCCCTCCCGCCGAGAGGGTCCAG 540
Db 629 AAAAAGAGTTACTGAAGCAGTATGGCTTTCCACGTGCGCCCTCCCGCCGAGAGGGTCCAG 688
Qy 541 GCTGCCAGGAGGACCGACAGCAGCTGTGCCCCCTGGAGCGGGTGATACAGAGAGATT 600
Db 689 GCTGCCAGGAGGACCGACAGCAGCTGTGCCCCCTGGAGCGGGTGATACAGAGAGATT 748
Qy 601 GCCATCTGAAGAAGCTGGACCACTGAATGTGTTCAAACTGATCGAGGTCTCTGATGAC 660
Db 749 GCCATCTGAAGAAGCTGGACCACTGAATGTGTTCAAACTGATCGAGGTCTCTGATGAC 808
Qy 661 CCAGCTGAGGACAACTCTATTGTGTTGACCTCTTGAGAAAGGGGCCCGTCAATGAA 720
Db 809 CCAGCTGAGGACAACTCTATTGTGTTGACCTCTTGAGAAAGGGGCCCGTCAATGAA 868
Qy 721 GTGCCCTGTGACAGCCCTTCTCGAGGAGCAAGCTCGCCCTCACTCGGGAGCTCATC 780
Db 869 GTGCCCTGTGACAGCCCTTCTCGAGGAGCAAGCTCGCCCTCACTCGGGAGCTCATC 928
Qy 781 CTGGCCCTCGAGTACTTGCATCTGCCAGAGATCGTCCACAGGACATCAAGCCATCCAAAC 840
Db 929 CTGGCCCTCGAGTACTTGCATCTGCCAGAGATCGTCCACAGGACATCAAGCCATCCAAAC 988
Qy 841 CTGTCTCTGGGGGATGATGGGCACTGGAAGATGCGCGACTTTGGCGTCAAGCAACGAGTTT 900
Db 989 CTGTCTCTGGGGGATGATGGGCACTGGAAGATGCGCGACTTTGGCGTCAAGCAACGAGTTT 1048
Qy 901 GAGGGAGAGCGCTCAGCTGTCCAGACGCGGGGAACCCAGCATTCATGGCCCCCGAG 960
Db 1049 GAGGGAGAGCGCTCAGCTGTCCAGACGCGGGGAACCCAGCATTCATGGCCCCCGAG 1108
Qy 961 GCCATTTCTGATTCGGGCGCAGAGCTTCAGTGGGAAGGCTTGGATGATGGGCCACTGGC 1020
Db 1109 GCCATTTCTGATTCGGGCGCAGAGCTTCAGTGGGAAGGCTTGGATGATGGGCCACTGGC 1168
Qy 1021 GTACGTTGTACTGCTTTGTCTATGGGAAGTGCCTCATTCGACGATTTGATCTCTGGCC 1080
Db 1169 GTACGTTGTACTGCTTTGTCTATGGGAAGTGCCTCATTCGACGATTTGATCTCTGGCC 1228
Qy 1081 CTCCACAGGAAGATCAAGAAATGAGCCGCTGGTGTCTCTGAGGAGCCAGAAATCAGCGAG 1140
Db 1229 CTCCACAGGAAGATCAAGAAATGAGCCGCTGGTGTCTCTGAGGAGCCAGAAATCAGCGAG 1288
Qy 1141 GAGCTCAAGGACCTGATCCTGAAGATGTTAGACAAGATCCCGAGACGAGAAATTTGGGGTG 1200
Db 1289 GAGCTCAAGGACCTGATCCTGAAGATGTTAGACAAGATCCCGAGACGAGAAATTTGGGGTG 1348
Qy 1201 CCAGACATCAAGTTGACCTTTGGGTGACCAAGAACGGGGAGGAGCCCTTCCTTCGGAG 1260
Db 1349 CCAGACATCAAGTTGACCTTTGGGTGACCAAGAACGGGGAGGAGCCCTTCCTTCGGAG 1408
Qy 1261 GAGGAGCACTGACGCTGTGAGGTGACAGAGGGGAGGTTTAAAGACTCAGTCAGGCTC 1320
Db 1409 GAGGAGCACTGACGCTGTGAGGTGACAGAGGGGAGGTTTAAAGACTCAGTCAGGCTC 1468
Qy 1321 ATCCCCAGCTGACCAACGCTGATCTCTGTGTAAGTCCATGCTGAGGAAGCGTTCTTTGGG 1380
Db 1469 ATCCCCAGCTGACCAACGCTGATCTCTGTGTAAGTCCATGCTGAGGAAGCGTTCTTTGGG 1528
Qy 1381 AACCCGTTTGGCCCCCAGCGAGGAGGAGCGGATCCATCTCTGCTCCAGGAACCTTA 1440
Db 1529 AACCCGTTTGGCCCCCAGCGAGGAGGAGCGGATCCATCTCTGCTCCAGGAACCTTA 1588

Qy 1441 CTGCTGAAGAGCGGTTTGGTGAAGGGGCAAGAGCCAGAGCTCCCGCGGTCCAGGAA 1500
Db 1589 CTGCTGAAGAGCGGTTTGGTGAAGGGGCAAGAGCCAGAGCTCCCGCGGTCCAGGAA 1648
Qy 1501 GACGAGGCTGCATCC 1515
Db 1649 GACGAGGCTGCATCC 1663
RESULT 7
AX746188 1937 bp DNA linear PAT 13-JUN-2003
LOCUS AX746188
DEFINITION Sequence 39 from Patent WO0208399.
ACCESSION AX746188
VERSION AX746188.1 GI:31746174
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS Yue, H., Khan, F. A., Gururajan, R., Hafalia, A. J., Chawla, N. K.,
Arvizu, C. S., Rankumar, J., Gandhi, A. R., Policky, J. L., Baughn, M. R.,
Tribouley, C. M., Bandman, O., Nguyen, D. B., Lu, Y., Burford, N., Lal, P.,
Ding, L., Yao, M. G., Elliott, V. S., Recipon, S. A., Kearney, L., Lu, D. A.,
Greenwald, S. R., Tang, Y. T., Xu, Y., Walsh, R. T., Gietzen, K. J.,
Yang, J., Jackson, J. L. and Thornton, M.
Human kinases
TITLE Patent: WO 0208399-A 39 31-JAN-2002;
JOURNAL Incyte Genomics, Inc. (US); Thornton, Michael (US)
FEATURES
source 1. 1937
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/note="Incyte ID No: 2190612CB1"

Query Match 99.7%; Score 1510.2; DB 6; Length 1937;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1512; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ATGAGGGGGGTCCAGCTGTCTGTGCCAGGATCTCGGGCAGAGCTGGTAGAACGGGTG 60
Db 65 ATGAGGGGGGTCCAGCTGTCTGTGCCAGGATCTCGGGCAGAGCTGGTAGAACGGGTG 124
Qy 61 GCAGCCATCGATGTGACTCACTTGGAGGAGGAGATGTGGCCAGAGCTTACTAGAAAC 120
Db 125 GCAGCCATCGATGTGACTCACTTGGAGGAGGAGATGTGGCCAGAGCTTACTAGAAAC 184
Qy 121 GGTGTGGACCCCCCACCACCGGCCAGAGCTGCCTCTGTGATCCCTGGCAGTACTTCAAGA 180
Db 185 GGTGTGGACCCCCCACCACCGGCCAGAGCTGCCTCTGTGATCCCTGGCAGTACTTCAAGA 244
Qy 181 CTGCTCCAGCCCCGCTTAGCTCTCAGCCAGGAGCTTCCCTACAGAGCGGCCAGCA 240
Db 245 CTGCTCCAGCCCCGCTTAGCTCTCAGCCAGGAGCTTCCCTACAGAGCGGCCAGCA 304
Qy 241 GGAAGCTATCTGGAGGCGCAGGCTTGGGCTTATGCCAGGGGCTTCCAGCCACATCTCC 300
Db 305 GGAAGCTATCTGGAGGCGCAGGCTTGGGCTTATGCCAGGGGCTTCCAGCCACATCTCC 364
Qy 301 CCGGGGCTTGGCGAGGCCCAACCATCGAGTCCCAACACGCTGGCCATCTCAGATCGAGAG 360
Db 365 CCGGGGCTTGGCGAGGCCCAACCATCGAGTCCCAACACGCTGGCCATCTCAGATCGAGAG 424
Qy 361 GACTGCTGACGTGAACCACTACAGCTGACAGTGAAGTGGCAAGGGTGCCTTACGGT 420
Db 425 GACTGCTGACGTGAACCACTACAGCTGACAGTGAAGTGGCAAGGGTGCCTTACGGT 484
Qy 421 GTGTGAGGCTGGCTTACAAGCAAGTGAAGACAGACACTATGCAATGAAAGTCTTTTCC 480

Db 485 GTGGTGAGGCTGGCCCTACAAACGAAAGTGAAGACAGACACTATATCAATGAAGTCTCTTCC 544
QY 481 AAAAGAAGTTACTGAAGCAGTATGGCTTTCCACGTCGCGCTCCCCCGAGAGGGTCCCGAG 540
Db 545 AAAAGAAGTTACTGAAGCAGTATGGCTTTCCACGTCGCGCTCCCCCGAGAGGGTCCCGAG 604
QY 541 GCTGCCAGGAGGACAGCAAGCAGCTGCTGCCCTGGAGCGGTGTACCAAGAGATT 600
Db 605 GCTGCCAGGAGGACAGCAAGCAGCTGCTGCCCTGGAGCGGTGTACCAAGAGATT 664
QY 601 GCATCTGAAGAGCTGACACAGCTGAATGTGGTCAAACTGATCAGGTCTCTGGATGAC 660
Db 665 GCATCTGAAGAGCTGACACAGCTGAATGTGGTCAAACTGATCAGGTCTCTGGATGAC 724
QY 661 CCAGCTGAGGACAACTTATTTGGTGTGTGACCTCTCAGAAAGGGCGCGTCATGGAA 720
Db 725 CCAGCTGAGGACAACTTATTTGGTGTGTGACCTCTCAGAAAGGGCGCGTCATGGAA 784
QY 721 GTGCCCTGTGACAAAGCCCTTCTCGAGGAGCAAGCTCGCTTACCTGCGGAGCGTCATC 780
Db 785 GTGCCCTGTGACAAAGCCCTTCTCGAGGAGCAAGCTCGCTTACCTGCGGAGCGTCATC 844
QY 781 CTGGGCTCGAGTACTGACCTGACAGAGAGTCTCCACAGGACATCAAGCCATCCAC 840
Db 845 CTGGGCTCGAGTACTGACCTGACAGAGAGTCTCCACAGGACATCAAGCCATCCAC 904
QY 841 CTGCTCTGGGGATGATGGGCACGTGAAGATCGCCGACTTTGGCGTCAGCAACCAAGTTT 900
Db 905 CTGCTCTGGGGATGATGGGCACGTGAAGATCGCCGACTTTGGCGTCAGCAACCAAGTTT 964
QY 901 GAGGGAAACGACGTGATGACGACGCGGGAAACCCAGCATTCATGAGCCCGGAG 960
Db 965 GAGGGAAACGACGTGATGACGACGCGGGAAACCCAGCATTCATGAGCCCGGAG 1024
QY 961 GCATTTCTGATTCGGGCGAGAGCTTCACTGGAGGCGCTTGATGATGGGCGCATGGC 1020
Db 1025 GCATTTCTGATTCGGGCGAGAGCTTCACTGGAGGCGCTTGATGATGGGCGCATGGC 1084
QY 1021 GTCACTTGTACTGCTTTGTCTATGGGAAGTCCCATTCATGACGATTTCACTCTGGCC 1080
Db 1085 GTCACTTGTACTGCTTTGTCTATGGGAAGTCCCGTTCATGACGATTTCACTCTGGCC 1144
QY 1081 CTCCACAGAAATCAAGATGAGCCCGTGGTTCCTTGAGAGCGCAGAAATCAGCGAG 1140
Db 1145 CTCCACAGAAATCAAGATGAGCCCGTGGTTCCTTGAGAGCGCAGAAATCAGCGAG 1204
QY 1141 GAGCTCAAGGACCTGATCCTGAAGATGTTAGACAAGAAATCCGAGACGAGATTCGGGTG 1200
Db 1205 GAGCTCAAGGACCTGATCCTGAAGATGTTAGACAAGAAATCCGAGACGAGATTCGGGTG 1264
QY 1201 CCAGACATCAAGTTGACCCCTTGGGTGACCAAGACGGGGAGAGCCCTTCTCTGGGAG 1260
Db 1265 CCAGACATCAAGTTGACCCCTTGGGTGACCAAGACGGGGAGAGCCCTTCTCTGGGAG 1324
QY 1261 GAGGACCTGACGCTGTGAGGTGACAGAGGGGAGGTTAAGAACTCAGTCAGGCTC 1320
Db 1325 GAGGACCTGACGCTGTGAGGTGACAGAGGGGAGGTTAAGAACTCAGTCAGGCTC 1384
QY 1321 ATCCCCAGCTGACCAAGCTGATCCTGTGGTGAAGTCCATCTCAGGAAGGTTCCCTTGGG 1380
Db 1385 ATCCCCAGCTGACCAAGCTGATCCTGTGGTGAAGTCCATCTCAGGAAGGTTCCCTTGGG 1444
QY 1381 AACCCGTTTGGCCCCAGCAGGAGGGAAGCGATCCATGCTCTCAGGAACCTTA 1440
Db 1445 AACCCGTTTGGCCCCAGCAGGAGGGAAGCGATCCATGCTCTCAGGAACCTTA 1504
QY 1441 CTGTTGAAGAGGTTTGGTGAAGGGGCAAGAGCCAGAGCTCCCGGCTCCAGGAA 1500
Db 1505 CTGTTGAAGAGGTTTGGTGAAGGGGCAAGAGCCAGAGCTCCCGGCTCCAGGAA 1564
QY 1501 GACGAGGCTGCATCC 1515
Db 1565 GACGAGGCTGCATCC 1579

RESULT 8
AX455763
LOCUS AX455763 1683 bp DNA linear PAT 06-JUL-2002
DEFINITION Sequence 9 from Patent WO0220800.
ACCESSION AX455763
VERSION AX455763.1 GI:21714790
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Meyers R.A. and Silos-Santiago I.
TITLE 1658, 14223, and 16002, novel human kinases and uses therefor
JOURNAL Patent: WO 0220800-A 9 14-MAR-2002;
MILLENNIUM PHARM INC (US)
FEATURES
Location/Qualifiers
1..1683
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 98.4%; Score 1490.6; DB 6; Length 1683;
Best Local Similarity 99.4%; Pred No 9.6e-312;
Matches 1496; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
QY 1 ATGAGGGGGGTCCAGCTGTCTGCTGCCAGGATCTCTGGGCGAGAGCTGGTAGAACGGGTG 60
Db 1 ATGAGGGGGGGTCCAGCTGTCTGCTGCCAGGATCTCTGGGCGAGAGCTGGTAGAACGGGTG 60
QY 61 GCAGCCATCGATGTCACTTCACTTGGAGGAGGAGATGTTGGGCCAGAGCCTACTAGAAC 120
Db 61 GCAGCCATCGATGTCACTTCACTTGGAGGAGGAGATGTTGGGCCAGAGCCTACTAGAAC 120
QY 121 GGTGTGGACCCGCCACCGGCCAGAGCTGCTCTGTATCCCTGGCAGTACTTCAAGA 180
Db 121 GGTGTGGACCCGCCACCGGCCAGAGCTGCTCTGTATCCCTGGCAGTACTTCAAGA 180
QY 181 CTGCTCCAGCCGGCGCTAGCTCTCAGCCAGGAAGCTTTCCTTACAGGAGCGGCAGCA 240
Db 181 CTGCTCCAGCCGGCGCTAGCTCTCAGCCAGGAAGCTTTCCTTACAGGAGCGGCAGCA 240
QY 241 GGAAGCTATCTGGAGCGGAGGCTGGGCTTTATGCCAGGGGGCTGCCAGCCATCTCC 300
Db 241 GGAAGCTATCTGGAGCGGAGGCTGGGCTTTATGCCAGGGGGCTGCCAGCCATCTCC 300
QY 301 CCCCAGGCTGGCGGAGGCCACCATCGAGTCCACACAGTGGCCATCTCAGATGCAGAG 360
Db 301 CCCCAGGCTGGCGGAGGCCACCATCGAGTCCACACAGTGGCCATCTCAGATGCAGAG 360
QY 361 GACTCGGTGAGCTGAACAGTCAAGCTGAGAGTGGCAAGGGTGCCTACGGT 420
Db 361 GACTCGGTGAGCTGAACAGTCAAGCTGAGAGTGGCAAGGGTGCCTACGGT 420
QY 421 GTGGTGAAGCTGGCTTACCAAGAAAGTGAACAGACACTATGCAATGAAGTCTTTCC 480
Db 421 GTGGTGAAGCTGGCTTACCAAGAAAGTGAACAGACACTATGCAATGAAGTCTTTCC 480
QY 481 AAAAAGAGTTACTGAAGCAGTATGGCTTTCCAGTCCGCTCCCGGAGGGTCCCGAG 540
Db 481 AAAAAGAGTTACTGAAGCAGTATGGCTTTCCAGTCCGCTCCCGGAGGGTCCCGAG 540
QY 541 GCTGCCAGGAGGACAGCCAGCAGCTGTGCTGCCCTGGAGCGGGTGTACCAAGAGATT 600
Db 541 GCTGCCAGGAGGACAGCCAGCAGCTGTGCTGCCCTGGAGCGGGTGTACCAAGAGATT 600
QY 601 GCCATCTTGAAGAGCTGGACCAAGTGTGTTCAAACTGATCGAGTCTCTGATGAC 660
Db 601 GCCATCTTGAAGAGCTGGACCAAGTGTGTTCAAACTGATCGAGTCTCTGATGAC 660

```
QY 661 CCAGCTGAGGACACCTCTATTGGTGTGTTGACCTCTGAGAAAGGGCCCGTCATGAA 720
Db 661 CCAGCTGAGGACACCTCTATTGGTGTGTTGACCTCTGAGAAAGGGCCCGTCATGAA 720
QY 721 GTGCCCTGTGACAAAGCCCTTCTCGGAGGAGCAAGCTGCGCTTACCTCGGGAGCGTCAIC 780
Db 721 GTGCCCTGTGACAAAGCCCTTCTCGGAGGAGCAAGCTGCGCTTACCTCGGGAGCGTCAIC 780
QY 781 GTGGCCCTCGAGTACTTGTCATGTCGACAGAGATGTCGCACAGGGACATCAAGCCATCCAAC 840
Db 781 GTGGCCCTCGAGTACTTGTCATGTCGACAGAGATGTCGCACAGGGACATCAAGCCATCCAAC 840
QY 841 CTGCTCTCGGGGATGATGGGACGTGAAGATCGCGACTTTGGCGTTCAGCAACACAGTTT 900
Db 841 CTGCTCTCGGGGATGATGGGACGTGAAGATCGCGACTTTGGCGTTCAGCAACACAGTTT 900
QY 901 GAGGGGAACGACGCTGAGCTGTCAGCAGCAGCGCGGGAACCCAGCATTCATGGCCCCGAG 960
Db 901 GAGGGGAACGACGCTGAGCTGTCAGCAGCAGCGCGGGAACCCAGCATTCATGGCCCCGAG 960
QY 961 GCCATTTCTGATTCGGCCAGAGCTTCTGATGGGAAGGCTTGGATGTATGGGCCACTGGC 1020
Db 961 GCCATTTCTGATTCGGCCAGAGCTTCTGATGGGAAGGCTTGGATGTATGGGCCACTGGC 1020
QY 1021 GTCAAGTTGTACTGCTTGTCTATGGGAAGTGCCTATTCATCGACGATTTTCATCTGSCC 1080
Db 1021 GTCAAGTTGTACTGCTTGTCTATGGGAAGTGCCTATTCATCGACGATTTTCATCTGSCC 1080
QY 1081 CTCCACAGGAAGATCAAGAATGAGCCCGTGTGTTCTCGAGGAGCCAGAAATCAGCGAG 1140
Db 1081 CTCCACAGGAAGATCAAGAATGAGCCCGTGTGTTCTCGAGGAGCCAGAAATCAGCGAG 1140
QY 1141 GAGCTCAAGGACCTGATCTCAAGATGTTAGCAAGAAATCCCGACAGAGAAATGGGGTG 1200
Db 1141 GAGCTCAAGGACCTGATCTCAAGATGTTAGCAAGAAATCCCGACAGAGAAATGGGGTG 1200
QY 1201 CCAGACATCAAGTTGACCCCTTGGGTGACCAAGAAAGGGGAGGACCCCTTCTCGGAG 1260
Db 1201 CCAGACATCAAGTTGACCCCTTGGGTGACCAAGAAAGGGGAGGACCCCTTCTCGGAG 1260
QY 1261 GAGGAGCACTGACGCTGGTGGAGGTGACAGGGGGAGGTTAAGAACTCAGTCAGGCTC 1320
Db 1261 GAGGAGCACTGACGCTGGTGGAGGTGACAGGGGGAGGTTAAGAACTCAGTCAGGCTC 1320
QY 1321 ATCCCGCTGACGACCGGTGATCTGTGAGTCCATGCTGAGGAGCGTTCCTTTGGG 1380
Db 1321 ATCCCGCTGACGACCGGTGATCTGTGAGTCCATGCTGAGGAGCGTTCCTTTGGG 1380
QY 1381 AACCCGTTTGAGCCCGGACGAGGAGGAGAGCGATCCATGTCTGTCTCCAGGAAACCTA 1440
Db 1381 AACCCGTTTGAGCCCGGACGAGGAGGAGAGCGATCCATGTCTGTCTCCAGGAAACCTA 1440
QY 1441 CTGTTGAAAGAGGGTTGTTGAGGGGGCAAGAGCCAGAGCTCCCGGGGTCCAGGAA 1500
Db 1441 CTGTTGAAAGAGGGTTGTTGAGGGGGCAAGAGCCAGAGCTCCCGGGGTCCAGGCT 1500
QY 1501 GAGCA 1505
Db 1501 TACCA 1505
```

```
RESULT 9
AX455761
LOCUS AX455761 2711 bp DNA linear PAT 06-JUL-2002
DEFINITION Sequence 7 from Patent WO0220800.
ACCESSION AX455761
VERSION AX455761.1 GI:21714788
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1
```

```
AUTHORS Meyers, R. A. and Silos-Santiago, I.
TITLE 16658, 14223, and 16002, novel human kinases and uses therefor
JOURNAL Patent: WO 0220800-A 7 14-MAR-2002;
MILLENNIUM PHARM INC (US)
FEATURES
source 1. .2711
/organism="Homo sapiens"
/db_type="unassigned DNA"
/db_xref="taxon:9606"
198..1883
/note="unnamed protein product"
Location/Qualifiers
1. .2711
/codon_start=1
/protein_id="CAD38244.1"
/db_xref="GI:21714789"
/db_xref="REMBL:CAD38244"
/translation="MEGPAVCCDDPAELVERVAADIVTHLEADGGPBTNRGVDP
PPRAASVIFPGTSRLLPAPSLARKLSQERPAQSYLEAQAPYATGPASHISPR
APRRPTTSEHVIAISDAEDCVLNQYKLSQIEIKGAYGVVIAVNESDRYAMKVL
KKLLKQYGFPRPPRPSOAGGPAKQLPLERVQETAILKKLQHNVVKLVLEVL
DDPAEDNLVLYFDLLRKGPVMEVPCDKPFSEQARLYLRDVILGLEVLCOKIVHRDI
KPSNLLGDDGHVKIADFGVSNQFEGNDAQLSSAGTAPAPWAPPAISDSQGSFSGKAL
DVWATGVTLVCFVYVKCFIDDFILALHRIKNEPVVFPPEPEISEEKDILLKMLDK
NPETRIGVDPDIKLHPWTKNEEPLSEEBHCSVVEVEEVKNSVLIPSWTTVILV
KSMLRKESFNPFPQARREBSMSAPENLIVKEGFGEGKSPGLPGVQAVYHNRDLP
LGPRPGSAATGIWRRGYPFAHLGVVAPELLSIQTLVWGLGLTGDRAVSGEF"
ORIGIN
Query Match 98.4%; Score 1490.6; DB 6; Length 2711;
Best Local Similarity 99.4%; Pred. No. 9.1e-312;
Matches 1496; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
QY 1 ATGAGAGGGGTCCAGCTGTCTGCTGCAGGATCCTCGGCAGAGCTGGTAGAACGGGTG 60
Db 198 ATGAGAGGGGTCCAGCTGTCTGCTGCAGGATCCTCGGCAGAGCTGGTAGAACGGGTG 257
QY 61 GCAGCCATCGATGTGACTCACTTGGAGGAGGACAGATGGTGGCCAGAGCCCTACTAGAAC 120
Db 258 GCAGCCATCGATGTGACTCACTTGGAGGAGGACAGATGGTGGCCAGAGCCCTACTAGAAC 317
QY 121 GGTGTGAGACCCCAACAGCGGCCAGAGCTGCCTCTGTGATCCCTGGCAGTACTTCAAGA 180
Db 318 GGTGTGAGACCCCAACAGCGGCCAGAGCTGCCTCTGTGATCCCTGGCAGTACTTCAAGA 377
QY 181 CTGCTCCAGCCCGGCTAGCCTCTCAGCCAGGAGGCTTTCCCTACAGGAGCGCCAGCA 240
Db 378 CTGCTCCAGCCCGGCTAGCCTCTCAGCCAGGAGGCTTTCCCTACAGGAGCGCCAGCA 437
QY 241 GGAAGCTATCTGAGGCGCAGGCTGGGCGCTTATGCCACGGGGCTTGCACGACCATCTCC 300
Db 438 GGAAGCTATCTGAGGCGCAGGCTGGGCGCTTATGCCACGGGGCTTGCACGACCATCTCC 497
QY 301 CCGCGGCTTGGCGAGGCGCCACCATCGAGTCCCAACAGCGTGGCCATCTCAGATGCGAG 360
Db 498 CCGCGGCTTGGCGAGGCGCCACCATCGAGTCCCAACAGCGTGGCCATCTCAGATGCGAG 557
QY 361 GACTGCGTGCAGCTGAACACAGTACAAGCTGCAGAGTGAAGATTGGCAAGGGTGCCTACGGT 420
Db 558 GACTGCGTGCAGCTGAACACAGTACAAGCTGCAGAGTGAAGATTGGCAAGGGTGCCTACGGT 617
QY 421 GTGTGAGGCTGGCCTACAACGAAAGTGAAGACAGACACTATATGCAATGAAGTCTCTTCC 480
Db 618 GTGTGAGGCTGGCCTACAACGAAAGTGAAGACAGACACTATATGCAATGAAGTCTCTTCC 677
QY 481 AAAAAGAGTTACTGAAGCAGTATGGCTTCCAGTCCCTCCCGAGAGGGTCCCGAG 540
Db 678 AAAAAGAGTTACTGAAGCAGTATGGCTTCCAGTCCCTCCCGAGAGGGTCCCGAG 737
QY 541 GCTCCCGAGGAGGAGCCAGCCAGCAGCTGTGTCCCTCTGGAGCGGGTGTATCAGAGAGATT 600
Db 738 GCTCCCGAGGAGGAGCCAGCCAGCAGCTGTGTCCCTCTGGAGCGGGTGTATCAGAGAGATT 797
QY 601 GCCATCTTGAAAGAGCTGGACCAGCTGAATGTGTGTAACACTGATCGAGGTCTCTGATCAC 660
```

16658, 14223, and 16002, novel human kinases and uses therefor
Patent: WO 0220800-A 7 14-MAR-2002;
MILLENNIUM PHARM INC (US)

FEATURES
source 1. .2711
/organism="Homo sapiens"
/db_type="unassigned DNA"
/db_xref="taxon:9606"
198..1883
/note="unnamed protein product"

Location/Qualifiers
1. .2711
/codon_start=1
/protein_id="CAD38244.1"
/db_xref="GI:21714789"
/db_xref="REMBL:CAD38244"
/translation="MEGPAVCCDDPAELVERVAADIVTHLEADGGPBTNRGVDP
PPRAASVIFPGTSRLLPAPSLARKLSQERPAQSYLEAQAPYATGPASHISPR
APRRPTTSEHVIAISDAEDCVLNQYKLSQIEIKGAYGVVIAVNESDRYAMKVL
KKLLKQYGFPRPPRPSOAGGPAKQLPLERVQETAILKKLQHNVVKLVLEVL
DDPAEDNLVLYFDLLRKGPVMEVPCDKPFSEQARLYLRDVILGLEVLCOKIVHRDI
KPSNLLGDDGHVKIADFGVSNQFEGNDAQLSSAGTAPAPWAPPAISDSQGSFSGKAL
DVWATGVTLVCFVYVKCFIDDFILALHRIKNEPVVFPPEPEISEEKDILLKMLDK
NPETRIGVDPDIKLHPWTKNEEPLSEEBHCSVVEVEEVKNSVLIPSWTTVILV
KSMLRKESFNPFPQARREBSMSAPENLIVKEGFGEGKSPGLPGVQAVYHNRDLP
LGPRPGSAATGIWRRGYPFAHLGVVAPELLSIQTLVWGLGLTGDRAVSGEF"

ORIGIN
Query Match 98.4%; Score 1490.6; DB 6; Length 2711;
Best Local Similarity 99.4%; Pred. No. 9.1e-312;
Matches 1496; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
QY 1 ATGAGAGGGGTCCAGCTGTCTGCTGCAGGATCCTCGGCAGAGCTGGTAGAACGGGTG 60
Db 198 ATGAGAGGGGTCCAGCTGTCTGCTGCAGGATCCTCGGCAGAGCTGGTAGAACGGGTG 257
QY 61 GCAGCCATCGATGTGACTCACTTGGAGGAGGACAGATGGTGGCCAGAGCCCTACTAGAAC 120
Db 258 GCAGCCATCGATGTGACTCACTTGGAGGAGGACAGATGGTGGCCAGAGCCCTACTAGAAC 317
QY 121 GGTGTGAGACCCCAACAGCGGCCAGAGCTGCCTCTGTGATCCCTGGCAGTACTTCAAGA 180
Db 318 GGTGTGAGACCCCAACAGCGGCCAGAGCTGCCTCTGTGATCCCTGGCAGTACTTCAAGA 377
QY 181 CTGCTCCAGCCCGGCTAGCCTCTCAGCCAGGAGGCTTTCCCTACAGGAGCGCCAGCA 240
Db 378 CTGCTCCAGCCCGGCTAGCCTCTCAGCCAGGAGGCTTTCCCTACAGGAGCGCCAGCA 437
QY 241 GGAAGCTATCTGAGGCGCAGGCTGGGCGCTTATGCCACGGGGCTTGCACGACCATCTCC 300
Db 438 GGAAGCTATCTGAGGCGCAGGCTGGGCGCTTATGCCACGGGGCTTGCACGACCATCTCC 497
QY 301 CCGCGGCTTGGCGAGGCGCCACCATCGAGTCCCAACAGCGTGGCCATCTCAGATGCGAG 360
Db 498 CCGCGGCTTGGCGAGGCGCCACCATCGAGTCCCAACAGCGTGGCCATCTCAGATGCGAG 557
QY 361 GACTGCGTGCAGCTGAACACAGTACAAGCTGCAGAGTGAAGATTGGCAAGGGTGCCTACGGT 420
Db 558 GACTGCGTGCAGCTGAACACAGTACAAGCTGCAGAGTGAAGATTGGCAAGGGTGCCTACGGT 617
QY 421 GTGTGAGGCTGGCCTACAACGAAAGTGAAGACAGACACTATATGCAATGAAGTCTCTTCC 480
Db 618 GTGTGAGGCTGGCCTACAACGAAAGTGAAGACAGACACTATATGCAATGAAGTCTCTTCC 677
QY 481 AAAAAGAGTTACTGAAGCAGTATGGCTTCCAGTCCCTCCCGAGAGGGTCCCGAG 540
Db 678 AAAAAGAGTTACTGAAGCAGTATGGCTTCCAGTCCCTCCCGAGAGGGTCCCGAG 737
QY 541 GCTCCCGAGGAGGAGCCAGCCAGCAGCTGTGTCCCTCTGGAGCGGGTGTATCAGAGAGATT 600
Db 738 GCTCCCGAGGAGGAGCCAGCCAGCAGCTGTGTCCCTCTGGAGCGGGTGTATCAGAGAGATT 797
QY 601 GCCATCTTGAAAGAGCTGGACCAGCTGAATGTGTGTAACACTGATCGAGGTCTCTGATCAC 660

```
Db 798 GCCATCTGGAAGAGCTGGACCACTGATGTGTCTAACTGATCGAGGTCTGGATGAC 857
Qy 661 CCAGCTGAGGAACACCTCTATTGGTGTGTTACCTCTCTGAGAAAGGGCCCGTCTATGGAA 720
Db 858 CCAGCTGAGGAACACCTCTATTGGTGTGTTACCTCTCTGAGAAAGGGCCCGTCTATGGAA 917
Qy 721 GTGCCCTGTGACAAAGCCCTTCTCGGAGGAGCAAGCTGCTCTACCTGCGGAGCGTCATC 780
Db 918 GTGCCCTGTGACAAAGCCCTTCTCGGAGGAGCAAGCTGCTCTACCTGCGGAGCGTCATC 977
Qy 781 CTGGCCCTCGAGTACTTGCCACTGACAGAGATCTCTCCACAGGACATCAAGCCATCCAA 840
Db 978 CTGGCCCTCGAGTACTTGCCACTGACAGAGATCTCTCCACAGGACATCAAGCCATCCAA 1037
Qy 841 CTGCTCTGGGGATGATGGGCACGTGAAGATCCCGACTTTGGCGTCAGCAACACAGTTT 900
Db 1038 CTGCTCTGGGGATGATGGGCACGTGAAGATCCCGACTTTGGCGTCAGCAACACAGTTT 1097
Qy 901 GAGGGGAACGACGCTCAGCTGTCCAGCACGCGGGGAACCCAGCAATTCATGGCCCCCGAG 960
Db 1098 GAGGGGAACGACGCTCAGCTGTCCAGCACGCGGGGAACCCAGCAATTCATGGCCCCCGAG 1157
Qy 961 GCCATTCTGATTCGGCCACAGAGCTTCAGTGGGAAGCCCTTGGATGATGGGCCACTGGC 1020
Db 1158 GCCATTCTGATTCGGCCACAGAGCTTCAGTGGGAAGCCCTTGGATGATGGGCCACTGGC 1217
Qy 1021 GTCACTGTGACTGTTTGTCTATGGGAAGTGCCCAATTCATGACGATTTTCATCTGGCC 1080
Db 1218 GTCACTGTGACTGTTTGTCTATGGGAAGTGCCCAATTCATGACGATTTTCATCTGGCC 1277
Qy 1081 CTCACAGGAAGATCAAGATGAGCCGCTGCTTCTGAGGAGCCAGCAAAATCAGCGAG 1140
Db 1278 CTCACAGGAAGATCAAGATGAGCCGCTGCTTCTGAGGAGCCAGCAAAATCAGCGAG 1337
Qy 1141 GAGCTCAAGGACCTGATCTCTGAAGATGTTAGACAAGAAATCCCGAGACGAAATGGGGTG 1200
Db 1338 GAGCTCAAGGACCTGATCTCTGAAGATGTTAGACAAGAAATCCCGAGACGAAATGGGGTG 1397
Qy 1201 CAGACATCAAGTTGACCTTGGGTGACCAAGAACGGGAGGAGCCCTTCTCTCGGAG 1260
Db 1398 CAGACATCAAGTTGACCTTGGGTGACCAAGAACGGGAGGAGCCCTTCTCTCGGAG 1457
Qy 1261 GAGGACATGACGCTGTGGAGGTGACAGAGGGGAGGTTAAGAACTCAGTCAGGCTC 1320
Db 1458 GAGGACATGACGCTGTGGAGGTGACAGAGGGGAGGTTAAGAACTCAGTCAGGCTC 1517
Qy 1321 ATCCCGAGCTGGACCAAGCTGATCTCTGGTGAAGTCCATGCTGAGGAAGCGTTCTTTGG 1380
Db 1518 ATCCCGAGCTGGACCAAGCTGATCTCTGGTGAAGTCCATGCTGAGGAAGCGTTCTTTGG 1577
Qy 1381 AACCCGTTTGAGCCCGCAGGACGAGGGAAGACGATCCATGCTGCTCCAGGAACCTTA 1440
Db 1578 AACCCGTTTGAGCCCGCAGGACGAGGGAAGACGATCCATGCTGCTCCAGGAACCTTA 1637
Qy 1441 CTGGTGAAGAAAGGTTTGGTGAAGGGGCAAGGCCAGAGCTCCCGCGCTCAGGAA 1500
Db 1638 CTGGTGAAGAAAGGTTTGGTGAAGGGGCAAGGCCAGAGCTCCCGCGCTCAGGAA 1697
Qy 1501 GACGA 1505
Db 1698 TACCA 1702
```

```
RESULT 10
AX166519
LOCUS AX166519 1542 bp DNA linear PAT 22-JUN-2001
DEFINITION Sequence 10 from Patent WO0138503.
ACCESSION AX166519
VERSION AX166519.1 GI:14546864
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

```
REFERENCE 1
AUTHORS Plowman,G.D., Whyte,D., Manning,G.S., Sudarsanam,S.S., Martinez,R.,
Flanagan,P. and Clary,D.S.
TITLE Novel human protein kinases and protein kinase-like enzymes
JOURNAL Patent: WO 0138503-A 10 31-MAY-2001;
Sugen, Inc. (US)
FEATURES
source 1..1542
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
```

```
Query Match 96.3%; Score 1458.8; DB 6; Length 1542;
Best Local Similarity 97.9%; Pred. No. 7.2e-305;
Matches 1510; Conservative 0; Mismatches 2; Indels 30; Gaps 2;

Qy 1 ATGAGAGGGGGTCCAGCTGTCTGCGCAGGATCTCTCGGCAGAGCTGTGTAGAACGGGTG 60
Db 1 ATGAGAGGGGGTCCAGCTGTCTGCTGCCAGATCTCTCGGCAGAGCTGTGTAGAACGGGTG 60
Qy 61 GAGGCCATCGATGTGACTCACTTTGGAGGAGGACATGGTGGCCCGACAGCCTACTAGAAAC 120
Db 61 GAGGCCATCGATGTGACTCACTTTGGAGGAGGACATGGTGGCCCGACAGCCTACTAGAAAC 120
Qy 121 GGTGTGAGCCCCCACCACGCGCCAGAGCTGCTCTGTGATCCCTGGCAGTACTTCAAGA 180
Db 121 GGTGTGAGCCCCCACCACGCGCCAGAGCTGCTCTGTGATCCCTGGCAGTACTTCAAGA 180
Qy 181 CTGCTCCCGCCCGGCTAGCCTCTCAGCCAGGAAGCTTTCCCTACAGGAGCGCCAGCA 240
Db 181 CTGCTCCCGCCCGGCTAGCCTCTCAGCCAGGAAGCTTTCCCTACAGGAGCGCCAGCA 240
Qy 241 GGAAGCTATCTGAGGCGCAGGCTGGGCCCTTATGCCACGCGGGCCTGCCAGCCACATCTCC 300
Db 241 GGAAGCTATCTGAGGCGCAGGCTGGGCCCTTATGCCACGCGGGCCTGCCAGCCACATCTCC 300
Qy 301 CCCCCGGCTGGCGAGGCGCCACCATCGAGTCCACACAGTGGCCCTGCCAGCCACATCTCC 360
Db 301 CCCCCGGCTGGCGAGGCGCCACCATCGAGTCCACACAGTGGCCCTGCCAGCCACATCTCC 360
Qy 361 GACTGCGTCAGCTGAACCAAGTACAAAGCTGCGAGTGAGATTGGCA----- 406
Db 361 GACTGCGTCAGCTGAACCAAGTACAAAGCTGCGAGTGAGATTGGCAAGGTGGGCTGACT 420
Qy 407 -----AGGGTGCCTAACGGTGTGCTGAGCTGGCCCTCAACCGAAAGTGAAGAC 453
Db 421 GATGCTATCTCAGGGTGCCTACGGTGTGCTGAGGCTGGCCTACAAACGAAAGTGAAGAC 480
Qy 454 AGACACTATGCAATGAAAGTCCCTTCCAAAAGAAAGTACTGAAGCAGTATGGCTTTCCA 513
Db 481 AGACACTATGCAATGAAAGTCCCTTCCAAAAGAAAGTACTGAAGCAGTATGGCTTTCCA 540
Qy 514 CCGTCCCTCCCGCAGAGGGTCCAGAGTGGCCAGGAGGAGGACCAAGCAGCTGCTG 573
Db 541 CCGTCCCTCCCGCAGAGGGTCCAGAGTGGCCAGGAGGAGGACCAAGCAGCTGCTG 600
Qy 574 CCGCTGGAGCGGGTGTACAGGAGATTGCCATCTCTGAAGAGCTGGACCACTGATGTG 633
Db 601 CCGCTGGAGCGGGTGTACAGGAGATTGCCATCTCTGAAGAGCTGGACCACTGATGTG 660
Qy 634 GTCAACTGATCGAGGTCTCGATGACCCAGCTGAGGACCACTTATTTGGTGTGAC 693
Db 661 GTCAACTGATCGAGGTCTCGATGACCCAGCTGAGGACCACTTATTTGGTGTGAC 717
Qy 694 CTCCTGAGAAAGGGGCCCGTCAATGGAAGTGCCTGTGCAAGCCCTTCTCGAGAGGACAA 753
Db 718 CTCCTGAGAAAGGGGCCCGTCAATGGAAGTGCCTGTGCAAGCCCTTCTCGAGAGGACAA 777
Qy 754 GCTGCTCTTACCTGCGGAGCGTCACTGCGGCTCGAGTACTTGCATGCCAGAGATC 813
Db 778 GCTGCTCTTACCTGCGGAGCGTCACTGCGGCTCGAGTACTTGCATGCCAGAGATC 837
```

QY	814	GTCCACAGGGGACATCAAGCCTCAACCTGCTCTCTGGGGATGATGGCCACCTGAAGATC	873
Db	838	GTCCACAGGGGACATCAAGCCTCAACCTGCTCTCTGGGGATGATGGCCACCTGAAGATC	897
QY	874	GCCGACTTTGGCGTCAGCAACAGTTTGAGGGGAACGACGCTCAGCTGTCCAGCACGGCG	933
Db	898	GCCGACTTTGGCGTCAGCAACAGTTTGAGGGGAACGACGCTCAGCTGTCCAGCACGGCG	957
QY	934	GGAAACCCAGCAATTCATGCGCCCGAGAGCCCAATTTCTGATTCGGGCCAGAGCTTCAGTGGG	993
Db	958	GGAAACCCAGCAATTCATGCGCCCGAGAGCCCAATTTCTGATTCGGGCCAGAGCTTCAGTGGG	1017
QY	994	AAGCCCTTGGATGTATGGGCCACTGGGCTCAGCTGTACTCTTTGCTATGGGAAGTGC	1053
Db	1018	AAGCCCTTGGATGTATGGGCCACTGGGCTCAGCTGTACTCTTTGCTATGGGAAGTGC	1077
QY	1054	CCATTTCAGACGATTTTCATCTCTGGCCCTCCACAGGAAGATCAAGATGAGCCCGTGGTG	1113
Db	1078	CCATTTCAGACGATTTTCATCTCTGGCCCTCCACAGGAAGATCAAGATGAGCCCGTGGTG	1137
QY	1114	TTTCTGAGGAGCCAGAAATCAGCGAGAGCTCAAGGACCTGATCCTGAAGATGTTAGAC	1173
Db	1138	TTTCTGAGGAGCCAGAAATCAGCGAGAGCTCAAGGACCTGATCCTGAAGATGTTAGAC	1197
QY	1174	AAGAAATCCGAGACGAGAAATGGGTCGACACATCAAGTTGCACCCCTGGGTGACCAAG	1233
Db	1198	AAGAAATCCGAGACGAGAAATGGGTCGACACATCAAGTTGCACCCCTGGGTGACCAAG	1257
QY	1234	AACGGGGAGGAGCCCTTCCTTCGGAGGAGGACACTGCAGCGTGGTGAGGATGACAGAG	1293
Db	1258	AACGGGGAGGAGCCCTTCCTTCGGAGGAGGACACTGCAGCGTGGTGAGGATGACAGAG	1317
QY	1294	GGGAGGTTAAGAACTCAGTACAGGCTCATCCCACTGGAACCAAGGATGATCCTGTGAAG	1353
Db	1318	GAGGAGTTAAGAACTCAGTACAGGCTCATCCCACTGGAACCAAGGATGATCCTGTGAAG	1377
QY	1354	TCCATGCTGAGGAGCGTTCTTTGGGAACCGTTTGAGCCCGACGAGGAGGAAAGAG	1413
Db	1378	TCCATGCTGAGGAGCGTTCTTTGGGAACCGTTTGAGCCCGACGAGGAGGAAAGAG	1437
QY	1414	CGATCCATGCTCTGCTCCAGGAACCTACTGTGTGAAGAGGTTTGTGAAGGGGGCAAG	1473
Db	1438	CGATCCATGCTCTGCTCCAGGAACCTACTGTGTGAAGAGGTTTGTGAAGGGGGCAAG	1497
QY	1474	AGCCAGAGCTCCCGGGCTCAGGAAGACGAGGCTGCATCC	1515
Db	1498	AGCCAGAGCTCCCGGGCTCAGGAAGACGAGGCTGCATCC	1539
RESULT 11			
AXI179641			
LOCUS			
Sequence 14 from Patent WO0146397.			
AXI179641			
ACCESSION			
AXI179641.1			
GI:15132054			
VERSION			
KEYWORDS			
Homo sapiens (human)			
SOURCE			
Homo sapiens			
ORGANISM			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
REFERENCE			
1			
Yang, J., Raughn, M.R., Burford, N., Au-Young, J., Lu, D.A., Reddy, R.,			
Yue, H., Yao, M.G., Lal, P. and Khan, F.A.			
Human kinases			
Patent: WO 0146397-A 14 28-JUN-2001;			
Incyte Genomics, Inc. (US)			
FEATURES			
Location/Qualifiers			
1..3501			
/organism="Homo sapiens"			
/mol_type="unassigned DNA"			
/db_xref="taxon:9606"			
/note="Incyte ID No: 2041716CB1"			
source			
ORIGIN			
Query Match 96.2%; Score 1457.2; DB 6; Length 3501;			
Best Local Similarity 97.9%; Pred. No. 1.5e-304;			
Matches 1509; Conservative 0; Mismatches 3; Indels 30; Gaps 2;			
QY	1	ATGAGGGGGGTCAGCTGTCTGTCAGGATCCTCGGCAGAGCTGGTAGAACGGGTG	60
Db	65	ATGAGGGGGGTCAGCTGTCTGTCAGGATCCTCGGCAGAGCTGGTAGAACGGGTG	124
QY	61	GCAGCCATCGATGTGATCTACTTGGAGGAGCAGATGGTGGCCAGAGCCCTACTAGAAC	120
Db	125	GCAGCCATCGATGTGATCTACTTGGAGGAGCAGATGGTGGCCAGAGCCCTACTAGAAC	184
QY	121	GCTGTGAGCCCCACACAGGCGCAGAGCTGCTCTGTGATCCCTGGCAGTACTTCAAGA	180
Db	185	GCTGTGAGCCCCACACAGGCGCAGAGCTGCTCTGTGATCCCTGGCAGTACTTCAAGA	244
QY	181	CTGCTCCAGCCCGGCTTAGCTCTCAAGCCAGGAGCTTCCCTACAGGAGCGGCGCAGCA	240
Db	245	CTGCTCCAGCCCGGCTTAGCTCTCAAGCCAGGAGCTTCCCTACAGGAGCGGCGCAGCA	304
QY	241	GGAGCTATCTGAGGCGCAGGCTGGCCCTTATGCCAGGGGCGCTGCCAGCCACATCTCC	300
Db	305	GGAGCTATCTGAGGCGCAGGCTGGCCCTTATGCCAGGGGCGCTGCCAGCCACATCTCC	364
QY	301	CCCGGGCTCGCGAGGCGCCACCATCGAGTCCACACACGTCGCCATCTCAGATCAGAG	360
Db	365	CCCGGGCTCGCGAGGCGCCACCATCGAGTCCACACACGTCGCCATCTCAGATCAGAG	424
QY	361	GACTGCGTGCAGCTGAACCACTCAAGCTGAGAGTGAATTGGCA-----	406
Db	425	GACTGCGTGCAGCTGAACCACTCAAGCTGAGAGTGAATTGGCAAGGTTGGGCGGTGACT	484
QY	407	-----AGGGTGCCTACGGTGTGGTGGCTGGCTCAACAGAAAGTCAAGAC	453
Db	485	GATGCTATCTGCGAGGTCCTACGGTGTGGTGGCTGGCTCAACAGAAAGTCAAGAC	544
QY	454	AGACACTATGCAATGAAAGTCTTTCCAAAAGAGTTACTGAAGCAGTATGGCTTTCCA	513
Db	545	AGACACTATGCAATGAAAGTCTTTCCAAAAGAGTTACTGAAGCAGTATGGCTTTCCA	604
QY	514	CGTGGCCCTCCCGAGAGGCTCCAGGCTGCCAGGAGGAGCAGCAGCAGCAGCTGCTG	573
Db	605	CGTGGCCCTCCCGAGAGGCTCCAGGCTGCCAGGAGGAGCAGCAGCAGCAGCTGCTG	664
QY	574	CCCTGGAGCGGCTTACAGAGATTCCTTGAAGAGCTGCAGCCAGCTGAATGTG	633
Db	665	CCCTGGAGCGGCTTACAGAGATTCCTTGAAGAGCTGCAGCCAGCTGAATGTG	724
QY	634	GTCAAACTGATCGAGGCTCTGATGACCCAGCTGAGGACCAACCTCTATTGGTGTGGAC	693
Db	725	GTCAAACTGATCGAGGCTCTGATGACCCAGCTGAGGACCAACCTCTATTGGTGTGGAC	781
QY	694	CTCCTGAGAAAGGGGCGCTGATGAAGTGCCTGTGCAAGCCCTTCTCGAGAGGACAA	753
Db	782	CTCCTGAGAAAGGGGCGCTGATGAAGTGCCTGTGCAAGCCCTTCTCGAGAGGACAA	841
QY	754	GCTCGCCTCTACTCTCGGACGCTCATCTGGGCTCTGAGTACTTGCACCTGCCAGAGATC	813
Db	842	GCTCGCCTCTACTCTCGGACGCTCATCTGGGCTCTGAGTACTTGCACCTGCCAGAGATC	901
QY	814	GTCCACAGGAGCATCAAGCCATCCAACTGTCTTGGGGATGATGGGCACTGAGATC	873
Db	902	GTCCACAGGAGCATCAAGCCATCCAACTGTCTTGGGGATGATGGGCACTGAGATC	961
QY	874	GCCGACTTTGGCTCAGCAACAGTTTGAGGGGAACGAGCTCAGCTGTCAGCAGCGGG	933
Db	962	GCCGACTTTGGCTCAGCAACAGTTTGAGGGGAACGAGCTCAGCTGTCAGCAGCGGG	1021
QY	934	GGAAACCCAGCAATTCATGGCCCGGAGGCGCATTTCTGATTCGGCCAGAGCTTCAAGTGG	993
Db	1022	GGAAACCCAGCAATTCATGGCCCGGAGGCGCATTTCTGATTCGGCCAGAGCTTCAAGTGG	1081

994 AAGCCCTTGATATGAGCCATGCGCTCACTGTTTCTGTTGATGGAAGTGC 1053
Db |||||
1082 AAGCCCTTGATATGAGCCATGCGCTCACTGTTTCTGTTGATGGAAGTGC 1141
Qy |||||
1054 CCATTTCATGACATTCATCTCTGCGCTCCACAGGAAGTCAAGATGAGCCGTGGT 1113
Db |||||
1142 CCGTTTCATGACATTCATCTCTGCGCTCCACAGGAAGTCAAGATGAGCCGTGGT 1201
Qy |||||
1114 TTTCCTGAGGAGCCAGAAATCAGGAGGAGCTCAAGGACCTGATCTGAGATGTTAGAC 1173
Db |||||
1202 TTTCCTGAGGAGCCAGAAATCAGGAGGAGCTCAAGGACCTGATCTGAGATGTTAGAC 1261
Qy |||||
1174 AGAATCCGAGACGAGAAATGCGGTGCCAGACATCAAGTTGACCCCTTGGGTGACCAAG 1233
Db |||||
1262 AAGAATCCGAGACGAGAAATGCGGTGCCAGACATCAAGTTGACCCCTTGGGTGACCAAG 1321
Qy |||||
1234 AACGGGAGGAGCCCTTCTTCGGAGGAGGACATCGACGTGTTGAGGTTGACAGAG 1293
Db |||||
1322 AACGGGAGGAGCCCTTCTTCGGAGGAGGACATCGACGTGTTGAGGTTGACAGAG 1381
Qy |||||
1294 GGGGAGGTTAAGAACTCAGTCAAGCTCATCCCGAGCTGGACACAGTGTCTCTGGTGAAG 1353
Db |||||
1382 GAGGAGTTAAGAACTCAGTCAAGCTCATCCCGAGCTGGACACAGTGTCTCTGGTGAAG 1441
Qy |||||
1354 TCATGCTGAGGAGGTTCTCTTGGGAAACCGGTTTGGAGCCCGACGAGGAGGAGAG 1413
Db |||||
1442 TCATGCTGAGGAGGTTCTCTTGGGAAACCGGTTTGGAGCCCGACGAGGAGGAGAG 1501
Qy |||||
1414 CGATCATGCTGCTCCAGGAACCTTCTGTTGAAAGAGGTTTGGTGAAGGGGCAAG 1473
Db |||||
1502 CGATCATGCTGCTCCAGGAACCTTCTGTTGAAAGAGGTTTGGTGAAGGGGCAAG 1561
Qy |||||
1474 AGCCAGAGCTCCCGGCGCTCCAGGAAGAGGAGGCTGTCATCC 1515
Db |||||
1562 AGCCAGAGCTCCCGGCGCTCCAGGAAGAGGAGGCTGTCATCC 1603

RESULT 12
BC031647
LOCUS
DEFINITION
Homo sapiens calcium/calmodulin-dependent protein kinase
alpha, transcript variant 3, mRNA (cDNA clone MGC:34095
IMAGE:5175128), complete cds.
BC031647
MGC.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 2469)
Straussberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
Diatchenko, L., Mariani, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stapleton, M., Soares, M.B., Bonaldo, M.P., Casavant, T.L.,
Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,
Carninci, P., Prange, C., Raha, S., Loquellano, N.A., Peters, G.J.,
Abrahamson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahney, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S.,
Sanchez, A., Whitting, M., Madan, A., Young, A.C., Shevchenko, Y.,
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzywinska, M.I., Skalska, U., Smallos, D.E.,
Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

22388257
12477932
2 (bases 1 to 2469)
Strausberg, R.
Direct Submission
Submitted (06-JUN-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgabbs@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: http://www.hgsc.bcm.tmc.edu/cdna/
Contact: amg@bcm.tmc.edu
Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Lousseged, H.,
Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Manavathi,
A.N., Gibbs, R.A.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAC Plate: 51 Row: 1 Column: 23
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 27437010.
Location/Qualifiers
1..2469
/organism="Homo sapiens"
/mol_type="RNA"
/db_xref="taxon:9606"
/clone="MGC:34095 IMAGE:5175128"
/tissue="Type="Brain, lung, Testis, adult, pooled whole"
/clone_lib="NIH MGC_115"
/lab_host="DH10B"
/note="Vector: PCMV-SPORT6"
1..2469
/gene="CAMKK1"
/note="synonyms: DKFZp761M0423, CAMKKA, MGC34095"
/db_xref="LocusID:84254"
83..1645
/codon_start=1
/product="calcium/calmodulin-dependent protein kinase 1
alpha, isoform b"
/protein_id="AAH31647.1"
/db_xref="GI:21594836"
/db_xref="LocusID:84254"
/translation="MEGGPVCDDPPRAELVVERVAADIVTHLEADGGPEPTRNGVDP
PPRAASVIFGTSRLLPAPPSLSARKLSQERPAQYLEAQGPYATGPAHLSR
AWRPTTSHHVAISDAEDCVLNQYKLSGKAGYGVRLAYNESDHYAMKVL
KKLKKQYFPRPPRSGSQAPKQLLERYQYIAIKKLDHNVVVKLEVL
DDPADNLIALQNAQIQLDSNTIAKPHSLPSEQDSGTSWAARSVOLLKRGPV
MEVCPDKPSEQARLYLRDVLGLYLHCOKIHRDIKPSNLLGDGDDHVKIADFGV
SNQEGNDQSLSTAGTAFNAPRAISDSGFSKGLDWATGVTLYCFVYKGPFI
DDFILAHRKIKNPVVPFPEPSEELKDLILKMLDKNPETRGVDPDKLHPWTKN
ERFDPSEEEHCVSVEVEEVKNSVRLIPSWTTVILVKSMLRKRSGNPFEPQARRE
GERSAPGNLLV"
464..1423
/note="S TKc; Region: Serine/Threonine protein kinases,
catalytic domain"
/db_xref="CDD:smart00220"

misc_feature
Query Match 87.2%; Score 1321.2; DB 9; Length 2469;
Best Local Similarity 92.0%; Pred. No. 3.9e-275;
Matches 1453; Conservative 0; Mismatches 13; Indels 114; Gaps 1;
Qy 1 ATGAGGGGGGTCCAGCTGTCTGTCAGAGATCTCTGGGCGAGAGCTGCTAGAACGGGTG 60
Db 83 ATGAGGGGGGTCCAGCTGTCTGTCAGAGATCTCTGGGCGAGAGCTGCTAGAACGGGTG 142

QY 61 GCAGCCATCGATGACTCACTTGGAGAGGCGAGATGGTGGCCCGCAGAGCCCTACTAGAAAC 120
Db 143 GCAGCCATCGATGACTCACTTGGAGAGGCGAGATGGTGGCCCGCAGAGCCCTACTAGAAAC 202
QY 121 GGTGTGAGACCCCCACACAGCGGCGCAGAGTGCCTCTGTGATCCCTTGGCAGTACTTCAAGA 180
Db 203 GGTGTGAGACCCCCACACAGCGGCGCAGAGTGCCTCTGTGATCCCTTGGCAGTACTTCAAGA 262
QY 181 CTGTCTCCAGCCCGGCTAGCTCTTCAAGCAGGAGAGTTCCTTACAGGAGCGGCGCAGCA 240
Db 263 CTGTCTCCAGCCCGGCTAGCTCTTCAAGCAGGAGAGTTCCTTACAGGAGCGGCGCAGCA 322
QY 241 GGAAGCTATCTGGAGGCGAGGCTGGGCTTATGCGACGGGCGCTGCCAGCCACATCTCC 300
Db 323 GGAAGCTATCTGGAGGCGAGGCTGGGCTTATGCGACGGGCGCTGCCAGCCACATCTCC 382
QY 301 CCCGGGCGCTGGCGAGGCGCCACCATCGAGTCCCAACACGTCGGCCATCTCAGATGCGAG 360
Db 383 CCCGGGCGCTGGCGAGGCGCCACCATCGAGTCCCAACACGTCGGCCATCTCAGATGCGAG 442
QY 361 GACTGCGTGCAGTGAACAGTACAGCTGCAGAGTGAATGGAAGGTCCTACGGT 420
Db 443 GACTGCGTGCAGTGAACAGTACAGCTGCAGAGTGAATGGAAGGTCCTACGGT 502
QY 421 GTGTGTAGGCTGGCTTACAAAGAGTGAAGACAGACACTATGCAATGAAAGTCTTTC 480
Db 503 GTGTGTAGGCTGGCTTACAAAGAGTGAAGACAGACACTATGCAATGAAAGTCTTTC 562
QY 481 AAAAAGAGTTACTGAAGCAGTATGGCTTTCCAGTCCGCTCCCGCCGAGAGGTCCTCCAG 540
Db 563 AAAAAGAGTTACTGAAGCAGTATGGCTTTCCAGTCCGCTCCCGCCGAGAGGTCCTCCAG 622
QY 541 GCTGCCAGGAGGACAGCCAGCAGCTGCTGCCCTGGAGCGGCTGACAGGAGTT 600
Db 623 GCTGCCAGGAGGACAGCCAGCAGCTGCTGCCCTGGAGCGGCTGACAGGAGTT 682
QY 601 GCCATCTCTGAAGAGCTGACACCTGAATGTGTGTCAAACTGATCGAGGTCCTTGATGAC 660
Db 683 GCCATCTCTGAAGAGCTGACACCTGAATGTGTGTCAAACTGATCGAGGTCCTTGATGAC 742
QY 661 CCAGCTAGGACAACTCTATTG----- 684
Db 743 CCAGCTAGGACAACTCTATTG----- 802
QY 685 ----- 684
Db 803 TCACAAATATGCCAAGCCCACTCTCCTGCTTCCTCTGAGCAGCAAGTGGATCC 862
QY 685 -----GTGTTTGACCTCTGAGAAAGGGGCGGTCATGGAAGTGCCC 726
Db 863 ACGTGGGCTGGCGCTCAGTGTGACCTCTGAGAAAGGGGCGGTCATGGAAGTGCCC 922
QY 727 TGTGCAAGACCTTCTCGAGAGGCAAGCTCGCTCTACCTGCGGAGCGTATCTCTGGG 786
Db 923 TGTGCAAGACCTTCTCGAGAGGCAAGCTCGCTCTACCTGCGGAGCGTATCTCTGGG 982
QY 787 CTCGAGTACTTGCCTGCGCAGAGATCGTCCACAGGGACATCAAGCCATCAACCTGTC 846
Db 983 CTCGAGTACTTGCCTGCGCAGAGATCGTCCACAGGGACATCAAGCCATCAACCTGTC 1042
QY 847 CTGGGGGATGATGGGACGTGAAGATCGCCAGCTTTGGCGTCGAGCAACAGTTTGAGGG 906
Db 1043 CTGGGGGATGATGGGACGTGAAGATCGCCAGCTTTGGCGTCAGCAACAGTTTGAGGG 1102
QY 907 AACGAGCTCAGCTGCAGACGCGGGAACCCAGCATTCATGCGCCCGCAGGCGATT 966
Db 1103 AACGAGCTCAGCTGCAGACGCGGGAACCCAGCATTCATGCGCCCGCAGGCGATT 1162
QY 967 TCTGATTCGGGCGAGAGCTTCACTGGGAGGCGCTTGATGTATGGGCGCACTGGCGTCAG 1026
Db 1163 TCTGATTCGGGCGAGAGCTTCACTGGGAGGCGCTTGATGTATGGGCGCACTGGCGTCAG 1222

QY 1027 TTGTACTGCTTTGTCTATGGGAAGTGCCCATTCATCGACGATTTTCATCTCGCCCTCCAC 1086
Db 1223 TTGTACTGCTTTGTCTATGGGAAGTGCCCGTTTCATCGACGATTTTCATCTCGCCCTCCAC 1282
QY 1087 AGGAAGATCAAGAAATGAGCCCGTGTGTTTCTTCTGAGGAGCCAGAAATCAGCGAGGAGCTC 1146
Db 1283 AGGAAGATCAAGAAATGAGCCCGTGTGTTTCTTCTGAGGAGCCAGAAATCAGCGAGGAGCTC 1342
QY 1147 AAGGACCTGATCTTGAAGATGTTAGACAAGNAATCCGAGAGGAGAAATTTGGGTGCCAGAC 1206
Db 1343 AAGGACCTGATCTTGAAGATGTTAGACAAGNAATCCGAGAGGAGAAATTTGGGTGCCAGAC 1402
QY 1207 ATCAAGTTGCAACCTTGGGTGACCAAGAACGGGAGGAGCCCTTCTCTTCGAGGAGGAG 1266
Db 1403 ATCAAGTTGCAACCTTGGGTGACCAAGAACGGGAGGAGCCCTTCTCTTCGAGGAGGAG 1462
QY 1267 CACTGCAGCGTGTGGAGGTGACAGAGGGGAGGTTTAAAGCTCAGTCAGTCAGCTCATCCCC 1326
Db 1463 CACTGCAGCGTGTGGAGGTGACAGAGGAGGAGGTTTAAAGCTCAGTCAGTCAGCTCATCCCC 1522
QY 1327 AGCTGGACCGGATGATCCTGTCGAGTCCATGCTGAGGAAGCGTTCCTTTGGGAACCCG 1386
Db 1523 AGCTGGACCGGATGATCCTGTCGAGTCCATGCTGAGGAAGCGTTCCTTTGGGAACCCG 1582
QY 1387 TTTCAGCCCCAGGACGAGGAGGAGCGATGCCATGCTGTCAGGAAACCTACTGTG 1446
Db 1583 TTTCAGCCCCAGGACGAGGAGGAGCGATGCCATGCTGTCAGGAAACCTACTGTG 1642
QY 1447 AAAGAAGGTTTGGTGAAG 1466
Db 1643 TAAGTACTGTGGCCAGGG 1662

RESULT 13
BC017529
LOCUS
DEFINITION Mus musculus calcium/calmodulin-dependent protein kinase kinase 1,
alpha, mRNA (cdna clone MGC:27706 IMAGE:4924656), complete cds.
ACCESSION BC017529
VERSION BC017529.1 GI:17028423
KEYWORDS MGC.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 3435)
AUTHORS Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
Scheetz,T.E., Brownstein,M.J., Ustin,T.B., Toshiyuki,S.,
Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
Abramson,R.D., Mullen,S.J., Bosak,S.A., McEwan,P.J.,
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Fahey,J., Helton,E., Kettman,M., Madan,A., Rodriguez,S.,
Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butterfield,Y.S., Krzywinski,M.I., Skalski,U., Smailus,D.E.,
Scherch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
JOURNAL
MEDLINE 22388257
PUBMED 12477932
REFERENCE 2 (bases 1 to 3435)
AUTHORS Strausberg,R.
TITLE Direct Submission
JOURNAL Submitted (19-NOV-2001) National Institutes of Health, Mammalian

Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: Jeffrey E. Green, M.D.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Genome Sequence Centre,

BC Cancer Agency, Vancouver, BC, Canada

info@bcgsc.bc.ca

Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield, Susanna Chan, Readman Chit, Chris Fjell, Erin Garland, Lee Guin, Letticia Hsiao, Martin Kryzanski, Reta Kutsche, Oliver Lee, Soo Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven Ness, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saeedi, Jacqueline Schein, Duane Smailus, Michael Smith, Lorraine Spence, Jeff Stott, Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy, George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAK Plate: 38 Row: p Column: 9
This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis, Similarity but not identity to protein.

FEATURES

source	Location/Qualifiers
1..3435	
/organism="Mus musculus"	
/mol_type="mRNA"	
/strain="FVB/N"	
/db_xref="taxon:10090"	
/clone="MGC:27706 IMAGE:4924656"	
/tissue_type="Salivary gland, 10 week old female mouse"	
/clone_lib="NCI CGAP_SG2"	
/lab_host="DH10B"	
/note="Vector: pCMV-SPORT6"	
1..3435	
/gene="Camk1"	
/notes="synonyms: Camk1, CamK1alpha"	
/db_xref="LocusID:55984"	
/db_xref="MGI:1891766"	
127..1644	
/codon_start=1	
/product="Camk1 protein"	
/protein_id="AAH17529.1"	
/db_xref="GI:17028424"	
/db_xref="LocusID:55984"	
/translation="MESGAVCCQDPRAELVDRVAALNVHALEADEGPEPARNGVDP PPRAASVIPSASRPTVPSLSARKFSQERFAGSLGQAVGYSTGPASHISPR SWRPETIESHRVAISDEPCVQLNQKLSGKGAIVGVRLAYNESEDRHYAMKVL KKKLLQGFPPRPFGSQATCGGPAKOLLPLRVYQETAILKKLDHVNVMKLVIL DDPAEDNLYLPDILLRKGPMVPCDPPPEOARLYRDILGLHYLHCKOIKVHRDI KPSNLLGDGHVKVADFGVSNQFENDAQLSAGTAPAPMAPEAISDSGQSPSGAL DVWATGYLYCGVCKCPIDYIILTLRHKINEAVFPPEPEVEDUKDLILRLDK NPETRIGSDIKLHWYKHGEELPFSHEEKHSVVEVEEVKNVRLIPSWITVILY KSMRLKRFGNFPFQPARRESRMSAPGSLMKEGCGCKSPFLPGVQDEAAS"	
508..1353	
/note="S_TKc; Region: Serine/Threonine protein kinases, catalytic domain"	
/db_xref="CDD:smart00220"	
ORIGIN	
Query Match	81.9%; Score 1241.4; DB 10; Length 3435;
Best Local Similarity	88.7%; Pred. No. 6.7e-258;
Matches 1344; Conservative	0; Mismatches 171; Indels 0; Gaps 0;
QY	1 ATGGAGGGGGTCCAGCTGTCTCTGCCAGGATCTCTCGGGCAGAGCTGTGTAGAACGGGTG 60
Db	127 ATGGAGAGTGGCCGACCGCTGTCTGCCAAGACCTCGTGCAGAGCTGTGTGGATCGGGTG 186
QY	61 GCAGCCATCGATGTGACTTCACCTTGAGGAGGCGAGATGGTGGCCCGAGAGCTACTAGAAAC 120

Db	187 GCGCCATCAATGTGGCCACTTGGAGAAGCAGATGAGGGTCCAGAGCCTGCCAGGAAT 246
QY	121 GGTGTGAGACCCCCACACACGGGGCAGAGCTCCCTTGTGTGTATCCCTGGCAGTACTTCAAGA 180
Db	247 GGTGTGATCTCTCCACCCCGGGCCAGAGCTGCTCCGTGTATCCCTGGCAGTGTCAAGA 306
QY	181 CTGCTCCAGCCCGGCTAGCTCTCAGCAGAGGAGCTTTCCTTACAGAGGGGGCAGCA 240
Db	307 CCCACTCAGTGGCCCCAGCTCTCGGTAGAAAGTTCTCCCTGCAGAAACGACAGCT 366
QY	241 GGAAGCTATCTGAGGGCCAGGCTGGGCTTATGCCACGGGGCTGCCAGACCATCTCC 300
Db	367 GGAAGCTGTCTGGGGCTCAAGTTGGCCCTTACTCTACAGGACCTGCAGTCACTCTCT 426
QY	301 CCCCGGCTCTGGCGAGGGCCACCATCGAGTCCACACCGTGGCCATCTCAGATGCGAG 360
Db	427 CCTCGGTCTCTGGCGAGACCCACCATCGAGTCCACCGTGTGGCCATCTCAGACACAG 486
QY	361 GACTGCTGAGCTGAACCACTACAGCTCAAGCTGCAGAGTCAGATTGCAAGGGTGCCTACGT 420
Db	487 GACTGTGTCAACTGAACCACTACAGCTGCAGAGTCAGATTGCAAGGGTGCCTATGGT 546
QY	421 GTGTGAGGCTGGCTTACAAACGAAAGTGAAGACAGACACTATGCAATGAAGTCTTTTC 480
Db	547 GTGTGAGGCTGGCTTACAAACGAAAGTGAAGACAGACACTATGCAATGAAGTCTTTTC 606
QY	481 AAAAAGATTACTGAAGCAGTATGGCTTTCACGTGGCCTCCCCCGAGAGGGTCCCG 540
Db	607 AAAAAGATTACTGAAGCAGTATGGCTTTCCTCGCGCTCTCCCCCAGAGGGTCCCAA 666
QY	541 GCTGCCCAGGGAGCAGCCACGACAGCTGCTGCCCTGGAGCGGTGTACAGAGATT 600
Db	667 GCTACCCAGGGAGGGCCAGCCAAACAGCTGCTGCCCTGGAGCGTGTGTATCAGGAGATT 726
QY	601 GCCATCTGAAGAAGCTGGACCAACGTAATGTGTGCTCAAACTGATCGAGGTCCTGGATGAC 660
Db	727 GCCATTCTAAAGAAGCTGGACCAACGTAATGTGTGATGATGATGATGATGATGATGAT 786
QY	661 CCAGCTGAGACAACTCTATTGGTGTGTGACCTCTCAGAAAGGGGGCCGCTCATGGAA 720
Db	787 CCGCTGAAAGCAACCTCTATTGGTGTTCGACCTCTCAGAAAGGGGGCCAGTCATGGA 846
QY	721 GTGCCCTGTGACAAGCCCTCTCGGAGGAGCAGCTCGCTCTACCTGGGGAGCTCATC 780
Db	847 GTGCCCTGCGACAGCCCTTCCAGAGGAGCAGCTCGCTCTACCTGGGGAGCATCATC 906
QY	781 CTGGGCTCTCAGTACTTGCACTGCCAGAGATCGTCCACAGGAGACATCAAGCCATCAAC 840
Db	907 CTGGGCTCTCAGTACTTGCACTGCCAGAGATGTGTCACAGGACATCAAGCCATCAAC 966
QY	841 CTGCTCTGGGGATGATGGGACAGTGAAGATCGCCGACTTTGGGTTCAGCAACCACTTT 900
Db	967 CTACTCTTGGGGACAGCGGATGTGAAGATCGCCGACTTTGGTGTCTAGTAACCACTTT 1026
QY	901 GAGGGAAACAGCTCAGCTGTCCAGCAGCGGGGAAACCCAGCATTATGCCCCCGAG 960
Db	1027 GAGGGAAATGATCTCAGCTGTCCAGTACGGCAGGAGCCCCGGCATTATGCCCCCGAG 1086
QY	961 GCCATTCTGATTCGGCCAGAGCTTCAGTGGAAAGGCTTGGATGTATGGGCCACTGGC 1020
Db	1087 GCCATTTCTGATTCGGCCAGAGCTTCAGTGGAAAGGCTTGGATGTATGGGCCACTGGG 1146
QY	1021 GTCAGTTGTACTGCTTGTCTATGGGAAGTGCCTCATTCAGCAGATTTTCATCTCTGGCC 1080
Db	1147 GTCAGCTATATTGCTTGTCTATGGGAAGTGCCTCATTCAGTACTTACTCTGACC 1206
QY	1081 CTCACAGGAAGATCAAGAATGAGCCCGTGTGTTTCTCTGAGGAGCCAGAAATCAGCAG 1140
Db	1207 CTGACCGTAAAGATCAAGAAATGAGCCCGTGTGTTTCTCTGAGGAGCCAGAGGTTCAGCG 1266
QY	1141 GAGCTCAAGGACCTGATCCTCGAAGATGTAGCAAGAAATCCCGACAGAGATTCGGGTG 1200

Db 1267 GACCTCAAGACCTGATCTCTGAGGATGCTAGCAAGAACTCTGAGACAGAAATGGGGTG 1326

Qy 1201 CCAGACATCAAGTTGACACCTTGGGTGACCAAGAACCGGGAGGAGCCCTTCTCTTGGAG 1260

Db 1327 TCAGATATCAAGTTACACCTTGGGTGACCAAGCATGGAGAGGAGCCCTTCTCTTCAAG 1386

Qy 1261 GAGGAGCTCAGCGTGGTGGAGGTGACAGAGGGGAGGTTAAGAACTCAGTCAGGCTC 1320

Db 1387 GAGGAGCACTGCAAGTGTGGTGGAGGTGACCGAGGAGAGAGGTCAAGAACTCAGTCAGGCTC 1446

Qy 1321 ATCCCGAGCTGACACCGTGTGATCTGCTGAAGTCCATGCTGAGGAAGCTTCCCTTTGGG 1380

Db 1447 ATCCCGAGCTGACACCTGATCTGCTGAAGTCCATGCTGAGGAAGCTTCCCTTTGGG 1506

Qy 1381 AACCCGTTTGAAGCCCCAGGACCGGAGGAGGAGCGATCCATGCTGCTGAGGAAGCTTCCCTTTGA 1440

Db 1507 AACCCATTGAGCCCCAAGCAGCGAGGAGGAAGAGATCCATGCTGCTGACCAAGGAGCTTA 1566

Qy 1441 CTGCTGAAGAGAGGTTTGGTGAAGGGGCGAGAGCCAGAGCTCCCGGGCTCCAGGAA 1500

Db 1567 CTGATGAAGAGAGATGCGGAGAGGTTGCAAAAGCCAGAGCTTCCCGGAGTCCAGGAG 1626

Qy 1501 GACGAGGCTGCATCC 1515

Db 1627 GATGAGGCTGCTGCC 1641

RESULT 14

LOCUS S83194

DEFINITION Ca2+/calmodulin-dependent protein kinase IV kinase isoform [rats, brain, mRNA, 3429 nt].

ACCESSION S83194

VERSION S83194.1 GI:1836160

KEYWORDS Rattus sp.

SOURCE Rattus sp.

ORGANISM Rattus sp.

REFERENCE 1. (bases 1 to 3429)

AUTHORS Okuno,S., Kitani,T. and Fujisawa,H.

TITLE Evidence for the existence of Ca2+/calmodulin-dependent protein kinase IV kinase isoforms in rat brain

JOURNAL J. Biochem. 119 (6), 1176-1181 (1996)

MEDLINE 96425004

PUBMED 8827455

REMARK GenBank staff at the National Library of Medicine created this entry [NCBI gibbsq 179633] from the original journal article.

FEATURES

source

1. .3429

Location/Qualifiers

/organism="Rattus sp."

/mol_type="mRNA"

/db_xref="taxon:10118"

1. .3429

/gene="Ca2+/calmodulin-dependent protein kinase IV kinase isoform, CaM-kinase kinase alpha"

107. .1624

/gene="Ca2+/calmodulin-dependent protein kinase IV kinase isoform, CaM-kinase kinase alpha"

/note="This sequence comes from Fig. 1; CaM-kinase kinase alpha"

/codon_start=1

/product="Ca2+/calmodulin-dependent protein kinase IV kinase isoform"

/protein_id="AAB46910.1"

/db_xref="GI:1836161"

/translation="MERSPAVCCDDPRAELVERVAALVAHLEAEEGPEPSNGVDP PPARASALFGASRFTVPRLSARKSLQRPAGSCLEAQGYSTGPASHMSR AWRPTIESHVAISDTECVQLNQYLQSEIGKAGVGVRLYAINREDRHYAMKVL KKKLLQYGFPRPPPGSQAPQGGAKQLPLERVYQETAILKKLDHVNVLIELV DDAEDNLVLFVOLLRKGPVMEVPCDKPFPEQARLYLRDIIILGLEYLHCQKIVHRDI

ORIGIN

Query Match 81.7%; Score 1238.2; DB 10; Length 3429;
Best Local Similarity 88.6%; Pred. No. 3.3e-257;
Matches 1342; Conservative 0; Mismatches 173; Indels 0; Gaps 0;

Qy 1 ATGAGAGGGGGTCCAGTGTCTGTGCGCAGAGTCTCTCGGCGAGAGCTGGTAGAAACGGGTG 60

Db 107 ATGAGAGCGATCCAGCGCTCTGTGCGCAAGACCTCGAGCGGAACCTGGTAGAGCGGGTG 166

Qy 61 GCAGGCATCGATGTGACTCACTTGGAGGAGCAGATGCTGGCCAGAGAGCTACTAGAAAC 120

Db 167 GCAGGCATCGATGTGGGCCACCTAGAGAGGAGAGAGGGGCCAGAGCTGCCAGCAAT 226

Qy 121 GGTGTGACCCGCCACACACCGGCGCAGAGCTGCCTCTGTGATCCCTGGCAGTACTTCAAGA 180

Db 227 GGTGTGACCCGCCACACCGGCGCAGAGCTGCCTCTGTGATCCCTGGCAGTACTTCAAGA 286

Qy 181 CTGTCTCCAGCCCGGCTTAGCTCTCAGCAGGAAGCTTTCCTCAGAGAGGGCCAGCA 240

Db 287 CCCACACAGTGCCTCCAGCGCTCTCTGTGAAAAGTTCTCCCTCAGGAAAGACGAGCG 346

Qy 241 GGAAGCTATCTGGAGGGCGCAGGCTGGGCTTATGCGACGGGGCTGCCAGCCACATCTCC 300

Db 347 GGAAGCTGTCTAGAGGCTCAGTTGGGCTTACTCTACAGGACCTGCCAGTCAATGTCT 406

Qy 301 CCCCAGGCTGTGGCGAGGCGCCACCATCGAGTCCCACCGCTGGCCATCTCAGATGCGAG 360

Db 407 CCTCGGGCTTGGCGGAGAGCCACCATCGAGTCCCACCATGTGGCCATCTCAGACACAG 466

Qy 361 GACTCGTGGTACGCTGAACAGTACAAAGCTCAGAGTGAGATTGGCAAGGGTGCCTACG 420

Db 467 GACTCGTGGTACGCTGAACAGTACAAAGCTCAGAGTGAGATTGGCAAGGGTGCCTATG 526

Qy 421 GTGGTGAAGCTGGCTGTACAAAGAAAGTGAAGACAGACATATGCAATGAAGTCTTTC 480

Db 527 GTGGTGAAGCTGGCTGTACAAAGAAAGTGAAGACAGACATATGCAATGAAGTCTTTC 586

Qy 481 AAAAAAGAGTTTACTGAAGCAGTATGGCTTTTCCAGTCGCGCCCTCCCGGAGAGGTCCT 540

Db 587 AAAAAAGAGTTTACTGAAGCAGTATGGCTTTTCCAGTCGCGCCCTCCCGGAGAGGTCCT 646

Qy 541 GCTGCCAGGAGGAGCCAGCAGCAGCTGCTGCCCTCGAGCGGTGTACCAAGAGATT 600

Db 647 GCTCCTCAGGGAGGCGCAGCAAGCAGCTGCTGCCCTCGAGCGGTGTATACCAAGAGATT 706

Qy 601 GCATCCTGAAGAAGCTGGACCAACGTAATGTGGTCAAACTGATCGAGTCTCTGGATGAC 660

Db 707 GCATTTTAAAGAACTGGACCAACGTAATGTGGTCAAACTGATCGAGTCTCTGGATGAT 766

Qy 661 CCAGCTGAGGACAACTCTATTTGGTGTGGTCTCTCAGAAAGGGGCCCTCATGGAA 720

Db 767 CCGTCTGAAGACAACTCTATTTGGTGTGGTCTCTCAGAAAGGGGACGAGTCATGGAA 826

Qy 721 GTGCCCTGTGACAAGCCCTTCTCGGAGGAGCAAGCTGCGCTCTACCTCGGGACGTCATC 780

Db 827 GTGCCCTGTGACAAGCCCTTCCAGAGGAGCAAGCTGCGCTCTACCTCGGGACATCATC 886

Qy 781 CTGGGCTCGAGTACTTGCATCTGCCAGAAAGATCGTCCAGAGGACATCAAGCCATCAAC 840

Db 887 CTGGGCTCGAGTACTTGCATCTGCCAGAAAGATGTGCACAGGGACATCAAGCGTCCAAT 946

Qy 841 CTGCTCTCGGGGATGATGGGACGTAAGATCGCGACTTTTGGCGTTCACCAACAGTTT 900

Db 947 CTGCTCTTGGGACGATGGGACGTAAGATCGCGACTTTTGGTGTCAACCAACAGTTT 1006

Qy 901 GAGGGGAAACACGCTCAGCTGTCCAGCAGCGCGGGAAACCCAGCAATTTCATGGCCCCGAG 960

Db 1007 GAGGGGAAATGATCTCAGCTGTCCAGTACCGCAGGAGGAGCCCGAGCATTCATGGCCCCGAG 1066


```
QY 961 GCCATTCTGATTCGGCCAGAGCTTCAGTGGGAAGCCCTTGGATGTATGGGCACCTGGC 1020
Db |||||
QY 1067 GCCATTCTGACACCGGCCAGAGCTTCAGTGGGAAGCCCTTGGATGTATGGGCACCTGGG 1126
Db |||||
QY 1021 GTCACGTTGTACTGCTTTGTTCTATGGAAGTGGCCATTCATCCAGCATTCCTCTGGCC 1080
Db |||||
QY 1127 GTTACACTGTATTGTTGTTCTAGCGGAAGTGCCCATTCATTGATGATGATCATCTGGCC 1186
Db |||||
QY 1081 CTCACAGGAAGATCAAGATGAGCCCGTGGTGTTCCTCGAGGAGCCAGAAATCAGCGAG 1140
Db |||||
QY 1187 CTCACAGGAAGATCAAGATGAGCCCGTGGTGTTCCTCGAGGAGCCAGAGTTCAGCGAG 1246
Db |||||
QY 1141 GAGCTCAAGACCTGATCCTGAAGATGTTAGACAAGAAATCCCGAGACGAGAAATTGGGGTG 1200
Db |||||
QY 1247 GAACTCAAGACCTGATCCTGAAGATGCTAGACAAGAAATCCTGAACAAGAAATTGGGGTG 1306
Db |||||
QY 1201 CAGACATCAAGTTGACCCCTTGGTGACCAAGAACGGGAGAGAGCCCTTCCTCTGGAG 1260
Db |||||
QY 1307 TCTGATATCAAGTTACACCCCTTGGTGACCAAGCATGGAGAGAGCCCTTCCTCTCAGAG 1366
Db |||||
QY 1261 GAGGACCTGACAGCGTGGTGGAGGTACAGAGGGGAGGTTAAGAACTCAGTCAGGCTC 1320
Db |||||
QY 1367 GAGGAACACTGCACTGTGGTAGAGTGACTGAGAGGAGGTGAAGAACTCAGTCAAGCTC 1426
Db |||||
QY 1321 ATCCCCAGCTGGACACCGGTGATCCTGGTGAAGTCCATGCTGAGGAAGCCTTCCTTTGGG 1380
Db |||||
QY 1427 ATCCCCAGCTGGACCACTGTGATCCTGGTCAAGTCTATGCTGAGAAAGCGTTCCTTTGA 1486
Db |||||
QY 1381 AACCCGTTTTCAGCCCGCAGGACCGGAGGAGAGAGGATCCATGCTGTCAGGAACCTTA 1440
Db |||||
QY 1487 AACCCATTTTTCAGCCCGCAGGACCGGAGGAGAGGATCTATGCTGCGCCAGGAACTTA 1546
Db |||||
QY 1441 CTGGTGAAGAAGAGGTTTGGTGAAGGGGCAAGAGCCAGAGTCCCGCGCTCCAGGAA 1500
Db |||||
QY 1547 CTGTTGAAGAGAGATGTGGAGAGGGGCAAGAGCCAGAGTTCCTCGAGTCCAGGAA 1606
Db |||||
QY 1501 GACGAGGCTGCATCC 1515
Db |||||
QY 1607 GATGAGGCTGCATCC 1621
Db |||||

RESULT 15
AB023658 3411 bp mRNA linear ROD 24-FEB-1999
LOCUS Rattus norvegicus mRNA for Ca/calmodulin-dependent protein kinase
DEFINITION kinase alpha, CaM-kinase kinase alpha, complete cds.
ACCESSION AB023658
VERSION AB023658.1 GI:4512333
KEYWORDS Ca/calmodulin-dependent protein kinase kinase alpha, CaM-kinase
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (sites)
AUTHORS Okuno,S., Kitani,T. and Fujisawa,H.
TITLE Evidence for the existence of Ca2+/calmodulin-dependent protein
kinase IV kinase isoforms in rat brain
JOURNAL J. Biochem. 119 (6), 1176-1181 (1996)
MEDLINE 96425004
PUBMED 8827455
REFERENCE 2 (bases 1 to 3411)
AUTHORS Fujisawa,H. and Okuno,S.
TITLE Direct Submission
JOURNAL Submitted (15-FEB-1999) Hitoshi Fujisawa, Asahikawa Medical
College, Department of Biochemistry; Nishikawa 4-5-3-11,
Asahikawa 078-8510, Japan (E-mail:Okuno@asahikawa-med.ac.jp,
Tel:+81-166-68-2340, Fax:+81-166-68-2349)
FEATURES
Location/Qualifiers
1..3411
/organism="Rattus norvegicus"
```

```
mol_type="mRNA"
db_xref="taxon.10116"
/tissue_type="brain"
/tissue_lib="brain CDNA"
CDS
107..1624
/codon_start=1
/product="Ca/calmodulin-dependent protein kinase kinase
alpha, CaM-kinase kinase alpha"
/protein_id="BAA75246.1"
/db_xref="GI:4512334"
/translation="MERSPAVCCDPRALVERVAIAISVAHLEEBEGPPEASNGVDP
PPRAASVTPGSASPTTPVPSLSARKFSLQERPAAGSCLEAQVGYSTPASGMSPPR
AKRPTIESHVAISDTECVQLQYKQSEIKGAYGVRLAYNREDRHYAMKVLIS
KKLLKQYGFPRRPPPGSOAQQPOGPAKOLLPLERVYQETAIILKLDHVNVLKLEVL
DIPAEMLYLVDLLKGPVMEVPCDKPPEQARLYLRDILLGLYLHCQKLVHEDI
KSNLLGGDDHVKIADFVGSQFEGNDAQLSSTAGTAPFAPPEAISLDTGQSFSGKAL
DWATATYFCVYKGFIDSEYILALHARKINEAVFPPEPEVSEBELKDLILMLDK
NPETRIGVSDIKLHPWVTKHGEPLFSEEHCSVVEVTEEVKNSVKLIPSWTTLIIV
KMLRKRSFQNPFPQARREERSMSPGNLLKKEGCGEGKSPELPGVQDEAAS"
3411
/note="18 a nucleotides"
polyA_site
ORIGIN
Query Match 81.6%; Score 1236.6; DB 10; Length 3411;
Best Local Similarity 88.5%; Pred. No. 7,4e-257;
Matches 1341; Conservative 0; Mismatches 174; Indels 0; Gaps 0;
QY 1 ATGGAGGGGGTCCAGCTCTCTGCTGCCAGGATCCTCGGCAGAGCTGTAGAACGGGTG 60
Db |||||
QY 107 ATGGAGCGGAGTCCAGCCCTGCTGCTCCAAGACCCCTCGAGCGGAACCTGTAGACGGGTG 166
Db |||||
QY 61 GCAGCCATCGATGTGACTCACTTGGAGGAGGAGATGGTGGCCCGACAGCCTACTAGAAAC 120
Db |||||
QY 167 GCAGCCATCAGTGTGGCCCACTAGAGAGGAGAGAGGCGCCAGCGCTGCACCAAT 226
Db |||||
QY 121 GGTGTGAGCCCCCACCACCGGGCCAGAGTGCCTCTGTGATCCCTGGCAGTACTTCAAGA 180
Db |||||
QY 227 GGTGTGAGCCCCCACCACCGGGCCAGAGTGCCTCTGTGATCCCTGGCAGTGCCTCAAGA 286
Db |||||
QY 181 CTGCTCCAGCCGGCTAGCCTCTCAGCCAGGAGCTTTCCTCAGAGAGCGGCCAGCA 240
Db |||||
QY 287 CCACACCAAGTGGTCCCGAGCCTCTCTGTAGAAAGTTCTCCTCGAGGAAACACAGCG 346
Db |||||
QY 241 GGAAGTATCTCGAGGCGCAGGCTGGGCGCTTATGCCACGGGGCGCTCCCGACCATCTCC 300
Db |||||
QY 347 GGAAGTGTCTAGAGCTCAGGTGGGCGCTTACTCTACAGGACCTGCCAGTCAATGTCT 406
Db |||||
QY 301 CCCCAGGCTGGGGAGGCCCCACCAFCGAGTCCCACACGTCGCCCATCTCAGATGACAGAG 360
Db |||||
QY 407 CCTCGGCTTGGCGGAGACCCACCATCGAGTCCCACCATGTGGCCATCTCAGACACAGAG 466
Db |||||
QY 361 GACTCGCTCAGCTGNAACCACTACAGCTGCAGATGAGATGGGCAAGGTGCTCTAGGTT 420
Db |||||
QY 467 GACTGTGTCAACTGAACCACTACAGCTGCAGATGAGATTTGGCAAGGTGCTCTATGTT 526
Db |||||
QY 421 GTGGTGAAGCTGGCCTACAAAGAAAGTGAAGACAGACACTATGCAATGAAAGTCCTTTCC 480
Db |||||
QY 527 GTGGTGAAGCTGGCCTACAAAGAAAGTGAAGACAGACACTATGCAATGAAAGTCCTTTCC 586
Db |||||
QY 481 AAAAAAGAGTTACTGAAGCAGTATGCGCTTTTCAACGTCGCCCTCCCCCGAGAGGGTCCAG 540
Db |||||
QY 587 AAAAAAGAGTTACTGAAGCAGTATGCGCTTTTCTCGCGTCTCCCCCGAGAGGGTCCCAA 646
Db |||||
QY 541 GCTGCCAGGAGGAGCCAGCAGCTGCTGCCCTCGAGCGGGTGTACACAGGAGATT 600
Db |||||
QY 647 GCTCCTCAGGAGGGCCAGCCAAACAGCTGCTGCCCTCGAGCGGTGTATACACAGGAGATT 706
Db |||||
QY 601 GCCATCTCTGAAGAAGCTGACCACTGGAATGTGGTCAAACTCATCGAGTCTCTGATGAC 660
Db |||||
QY 707 GCATTTTAAAGAAACTGGACCACTGATGTAGTCAAGTGTCAAGTGTCTCTGATGAT 766
Db |||||
QY 661 CCAGCTGAGGACCAACTCTATTGTTGTTGATCTCTCCTGAGAAAGGGCCCGCTCATGAA 720
Db |||||
```

Db 767 CCTGCTGAAGACAATCTCTATTTTGGTGTGTGACCTCTGAGAAAGGACCAGTCATGGAA 826
 QY 721 GTGCCCTGTGACAAGCCCTTCTGGAGGAGCAAGCTCGCTCTACTGTGGGGACGTATC 780
 Db 827 GTGCCCTGTGACAAGCCCTTCCAGAGGAGCAAGCTCGCTCTACTGTGGGGACATCATC 886
 QY 781 CTGGSCCTTCAGTACTTCCACTGCCAGAGATCGTCCACAGGCATCAAGCCATCCAAC 840
 Db 887 CTGGSCCTTCAGTACTTCCACTGCCAGAGATTTGTCACAGGCATCAAGCCGTCCAAT 946
 QY 841 CTGCTCTCTGGGGATGATGGGCACTGAGATCGCGACTTTGGCGTCAAGCAACCAAGTTT 900
 Db 947 CTGCTCTCTGGGGACATGGGCACTGAGATCGCGACTTTGGTGTGAGCAACCAAGTTT 1006
 QY 901 GAGGGGAAGCGCTCAGCTGTCCAGACGGGGGAAACCCAGCAATTCATGGCCCCCGAG 960
 Db 1007 GAGGGGAATGATGCTCAGCTGTCCAGTACGGCAGGACCCACAGCAATTCATGGCCCCCGAG 1066
 QY 961 GCCATTTCTGATTCGGGCCAGAGCTTCAGTGGGAAGGCTTGGATGTATGGGCCACTGGC 1020
 Db 1067 GCCATCTCTGACACCGGCCAGAGCTTCAGTGGGAAGGCTTGGATGTATGGGCCACTGGG 1126
 QY 1021 GTCACTGTACTGTCTTTGCTATGGGAAGTCCCAATTCATCAGCAATTTTCATCTGGCC 1080
 Db 1127 GTTACACTGTATTGTTTGTCTACGGGAAGTCCCAATTCATTCATGATGATGATCATCTGGCC 1186
 QY 1081 CTCCACAGAGATCAAGATCAGCCCGTGGTGTTCCTGAGGAGCCAGAAATCAGCGAG 1140
 Db 1187 CTGCAAGAAAGATCAAGAAATGAGGCGCTGGTGTTCCTGAGGAGCCGAGGTCAAGCGAG 1246
 QY 1141 GAGCTCAAGGACCTGATCTGAAGATGTTAGACAAGATCCCGAGACGAGAAATTCGGGTG 1200
 Db 1247 GAACTCAAGACCTGATCTGAAGATGCTAGACAAGATCCCTGAACAAGAAATTCGGGTG 1306
 QY 1201 CCAGACATCAAGTTGCACCTTTGGGTGACCAAGAACGGGGAGGAGCCCTTCCTTCGGAG 1260
 Db 1307 TCTGATATCAAGTTACACCTTTGGGTGACCAAGCATGGAGAGGAGCCCTTCCTTCAGAG 1366
 QY 1261 GAGGAGCACTGAGCGTGTGGAGGTGACAGAGGGGGAGTTAAGAACTCAGTCAGGCTC 1320
 Db 1367 GAGGAACACTGCAGTGTGGTAGAGGTGACTGAGGAGGAGGTGAAGAACTCAGTCAGGCTC 1426
 QY 1321 ATCCCCAGCTGGACCAACGCTGATCCTGGTGAAGTCCATGCTGAGGAAGGTTCCCTTTGGG 1380
 Db 1427 ATCCCCAGCTGGACCACTGTGATCCTGGTCAAGTCTATGCTGAGAAAGGTTCCCTTTGA 1486
 QY 1381 AACCCGTTTGAGCCCGCAGGCAAGGAGGAGGAGGATCCATGCTGCTCCAGGAAACCTTA 1440
 Db 1487 AACCCATTTGAGCCCGCAGGCAAGGAGGAGGAGATCTATGCTGCGCCAGGAAACTTA 1546
 QY 1441 CTGGTGAAGAGAGGTTTGGTGAAGGGGCAAGAGCCCAAGAGCTCCCGGCGTCCAGGAA 1500
 Db 1547 CTGTTGAAGAGAGGATGTGGAGAAAGGGGCAAGAGCTTCCTCGGAGTCCAGGAA 1606
 QY 1501 GACGAGGCTGCATCC 1515
 Db 1607 GATGAGGCTGCATCC 1621

Search completed: July 25, 2004, 09:54:21
 Job time : 9539 secs